



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE CIENCIAS

BIOMEDICINA

Análisis de polimorfismos de nucleótido único implicados en cambios sinónimos y no sinónimos en moléculas de la inmunidad innata en relación con la respuesta al tratamiento en pacientes infectados con Hepatitis C.

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRA EN CIENCIAS BIOLÓGICAS

PRESENTA:

LBD. ALI BERENICE POSADA REYES

TUTOR PRINCIPAL DE TESIS: DR. SALVADOR FONSECA CORONADO

FACULTAD DE CIENCIAS, UNAM.

COMITÉ TUTOR: DRA. DIAZ BARRIGA ARCEO SANDRA

FACULTAD DE CIENCIAS, UNAM.

DR. EDGAR ENRIQUE SEVILLA REYES

FACULTAD DE MEDICINA, UNAM.

CIUDAD DE MÉXICO, ENERO, 2017.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE CIENCIAS

BIOMEDICINA

Análisis de polimorfismos de nucleótido único implicados en cambios sinónimos y no sinónimos en moléculas de la inmunidad innata en relación con la respuesta al tratamiento en pacientes infectados con Hepatitis C.

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRA EN CIENCIAS BIOLÓGICAS

PRESENTA:

LBD. ALI BERENICE POSADA REYES

TUTOR PRINCIPAL DE TESIS: DR. SALVADOR FONSECA CORONADO
FACULTAD DE CIENCIAS, UNAM.

COMITÉ TUTOR: DRA. DIAZ BARRIGA ARCEO SANDRA
FACULTAD DE CIENCIAS, UNAM.

DR. EDGAR ENRIQUE SEVILLA REYES
FACULTAD DE MEDICINA, UNAM.

CIUDAD UNIVERSITARIA, CD. MX., ENERO, 2017.

POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE CIENCIAS
DIVISIÓN DE ESTUDIOS DE POSGRADO

OFICIO FCIE/DEP/026/2017

ASUNTO: Oficio de Jurado

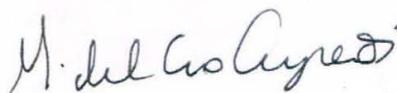
Lic. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día **17 de octubre de 2016** se aprobó el siguiente jurado para el examen de grado de **MAESTRA EN CIENCIAS BIOLÓGICAS** en el campo de conocimiento de **Biomedicina** del (la) alumno (a) **POSADA REYES ALI BERENICE** con número de cuenta **410009033** con la tesis titulada "**Análisis de polimorfismos de nucleótido único implicados en cambios sinónimos y no sinónimos en moléculas de la inmunidad innata en relación con la respuesta al tratamiento en pacientes infectados con Hepatitis C**", realizada bajo la dirección del (la) **DR. SALVADOR FONSECA CORONADO**:

Presidente: DR. RAMÓN MAURICIO CORAL VÁZQUEZ
Vocal: DR. ALEJANDRO ESCOBAR GUTIÉRREZ
Secretario: DR. EDGAR ENRIQUE SEVILLA REYES
Suplente: DRA. YALBI ITZEL BALDERAS MARTÍNEZ
Suplente: DRA. LETICIA MORENO FIERROS

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPÍRITU"
Ciudad Universitaria, Cd. Mx., a 13 de enero de 2017



DRA. MARÍA DEL CORO ARIZMENDI ARRIAGA
COORDINADORA DEL PROGRAMA



MCAA/MJFM/ASR/ipp

AGRADECIMIENTOS

Primeramente, al Posgrado en Ciencias Biológicas, UNAM por brindarme la oportunidad de estar dentro programa de Maestría y la aportación de diversos apoyos, que encaminaron mi proyecto de investigación.

Agradezco al Consejo Nacional de Ciencia y Tecnología por el apoyo recibido a través de la beca CONACYT 631936, para la realización de mis estudios de Maestría y su respaldo a los proyectos Salud-2012-C01-181585 “Análisis de la Dinámica y Evolución Intrahospedero de las Variantes Virales Resistentes a los Fármacos de Nueva Generación contra el Virus de la Hepatitis C” y Salud 219508 “Dinámica de Evolución y compartimentación de variantes virales resistentes a los antivirales de acción directa y del perfil de la respuesta inmunológica intrahepática y periférica en pacientes con infección crónica con virus de la hepatitis C”, colaboración bilateral con el Consejo Nacional de Ciencia y Tecnología en México y el Ministerio de Ciencia y Tecnología en Argentina.

Agradezco a mi Tutor principal Dr. Salvador Fonseca Coronado y a los miembros de mi Comité Tutor Dra. Sandra Díaz Barriga Arceo y Dr. Edgar Enrique Sevilla Reyes, por todo el apoyo, paciencia, disposición y dedicación brindada durante la realización del proyecto que contribuyo a dirigir mi formación académica.

AGRADECIMIENTOS A TITULO PERSONAL

- Al Dr. Ramón Mauricio Coral Vázquez, Dr. Alejandro Escobar Gutiérrez, Dra. Leticia Moreno Fierros y Dra. Yalbi Itzel Balderas Martínez, por toda su paciencia y contribuciones al presente trabajo.
- A la M en C. Sara Sixtos y el Dr. Francisco Sánchez del Departamento de Gastroenterología del Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán por todo el apoyo y confianza que brindaron al proyecto.
- Al Dr. Gilberto Vaughan por toda su colaboración en el proyecto de investigación, sobre todo por el sinfín de puertas abiertas y recorridas durante el mismo.
- A mis amigas Saraí Quiroz Cruz y Jerusalén Islas Arenas, sin ustedes chicas no hubiera sido posible, agradezco de todo corazón el apoyo incondicional que me proporcionaron y les deseo que el éxito las acompañe siempre.
- A todos y cada uno de los investigadores, maestros, compañeros del posgrado, que respondieron todas y cada una de mis dudas, me brindaron su apoyo y buscaron el mejor consejo para compartirlo.

DEDICATORIA

Este trabajo es dedicado, a las personas más importantes en mi vida, mi familia: Delfino, Lupita, Josue Arturo y Ofelia, gracias por todo su amor, paciencia, apoyo y todos los kilómetros recorridos durante este viaje en beneficio de la de la ciencia, esto es de ustedes.

ÍNDICE GENERAL

ÍNDICE DE FIGURAS.....	i
ÍNDICE DE TABLAS.....	ii
ABREVIATURAS.....	ii
RESUMEN.....	5
ABSTRACT.....	7
1. Introducción.....	9
1.1 Hepatitis C.....	9
1.2 Agente Etiológico.....	10
1.3 Heterogeneidad y variabilidad genética del HCV.....	13
1.4 Transmisión del HCV.....	15
1.5 Progresión de la enfermedad por HCV.....	16
1.6 Epidemiología.....	17
1.7 Tratamiento.....	18
1.8 Tipos de Respuesta al tratamiento.....	20
1.9 Factores que intervienen en la respuesta al tratamiento.....	22
1.9.1 Polimorfismos de nucleótido único.....	23
1.9.2 Técnicas de identificación de SNPs.....	25
1.9.3 Secuenciación.....	25
1.9.5 Paquete Bioinformático STADEN.....	27
1.9.6 Secuenciación Masiva en paralelo para identificar variantes virales intrahospedero de HCV.....	28
1.9.7 Herramientas bioinformáticas para el análisis de datos por secuenciación masiva en paralelo.....	30
2. Objetivos.....	32
2.1 Objetivo General.....	32
3. Justificación.....	32
4. Metodología.....	33
4.1 Diseño de oligonucleótidos específicos para la identificación de las moléculas TICAM1 y NOD1.....	34
4.2 Población de Estudio.....	36
4.3 Extracción de DNA y Genotipificación de SNPs en las regiones que codifican para las moléculas TICAM1 y NOD1.....	36
4.4 Análisis Bioinformático de la secuenciación de SNPs y predicción in silico.....	37
4.5 Análisis Estadístico de la Secuenciación de SNPs.....	37

4.4	Extracción de RNA Viral y Secuenciación de la Región HVR1 por NGS.....	37
4.6	Análisis Bioinformático de la región HVR1.....	38
5.	Resultados	38
5.1	Descripción del tipo de respuesta al tratamiento en pacientes infectados con HCV sometidos a terapia triple.....	38
5.2	Identificación de SNPs presentes en las moléculas de la Inmunidad Innata TICAM1 y NOD1.	39
5.2.1	Descripción de SNPs en la molécula TICAM1.....	39
5.2.3	Modelo de Correlación entre los SNPs en TICAM1 y NOD1 en los pacientes con HCV.....	46
5.3	Caracterización de la región HVR1 en pacientes con HCV sometidos a terapia triple.....	47
6	Discusión.....	50
7	Conclusiones.....	57
8.	Prespectivas.....	58
9.	Literatura Citada.	58

ÍNDICE DE FIGURAS

Figura 1. Estructura del HCV.....	10
Figura 2. Organización genética y procesamiento de la poliproteína del HCV.....	11
Figura 3. Genoma del virus y variabilidad nucleotídica del HCV.....	15
Figura 4. Distribución de los genotipos de HCV en la República Mexicana.....	17
Figura 5. Gráfica de la carga viral en pacientes respondedores en HCV.....	22
Figura 6. Gráfica de la carga viral en pacientes no respondedores en HCV.....	22
Figura 7. Esquema que representa un SNP.....	23
Figura 8. Diagrama general de trabajo en la técnica de secuenciación automatizada de Sanger.....	26
Figura 9. Diagrama general de trabajo del paquete Staden,.....	28
Figura 10. Diagrama general de trabajo de la plataforma 454 de ROCHE.....	30
Figura 11. Estrategia general experimental de trabajo. Se observa la población de estudio 6 pacientes infectados con HCV sometidos a terapia triple.....	33
Figura 12. Estrategia del diseño de oligonucleótidos para la molécula TICAM1.....	34
Figura 13. Estrategia del diseño de oligonucleótidos para la molécula NOD1.....	35
Figura 14. Cargas virales (\log_{10}) de los 4 pacientes Respondedores.....	39
Figura 15. Cargas virales (\log_{10}) de los 2 pacientes No Respondedores.....	39
Figura 16. Imagen representativa de la cobertura de la molécula TICAM1.....	40
Figura 17. Electroferogramas del primer amplicón de la molécula TRIF.....	40
Figura 18. Electroferogramas del primer amplicón de la molécula TRIF.....	41
Figura 19. Ensamble de secuencias del amplicón 5 de TRIF.....	41
Figura 20. Electroferogramas del quinto amplicón de la molécula TRIF.....	41
Figura 21. Electroferogramas del quinto amplicón de la molécula TRIF.....	42
Figura 22. Predicción in silico de la molécula TICAM1.....	42
Figura 23. Imagen representativa del diseño de primers para la secuenciación de la molécula NOD1.....	43
Figura 24. Electroferogramas del tercer amplicón de la molécula NOD1.....	43
Figura 25. Electroferogramas del tercer amplicón de la molécula NOD1.....	44
Figura 26. Ensamble de secuencias del amplicón 4 de NOD1.....	44
Figura 27. Electroferogramas del cuarto amplicón de la molécula de NOD1.....	45
Figura 28. Electroferogramas del cuarto amplicón de la molécula de NOD1.....	45
Figura 29. Predicción in silico de la molécula NOD1.....	45
Figura 30. Prueba de correlación Tau B Kendall.....	46
Figura 31. Promedios de las distancias nucleotídicas en la región HVR1 en 4 pacientes infectados con HCV sometidos a terapia triple.....	47
Figura 32. Análisis filogenético por el método de Maximum likelihood.....	49
Figura 33. Análisis filogenético por el método de Maximum likelihood.....	49
Figura 34. Vía de Señalización mediada por TLR3.....	51
Figura 35. Vía de Señalización de la molécula NOD1.....	53
Figura 36. Descripción general de la respuesta inmune innata antiviral para HCV.....	55

ÍNDICE DE TABLAS

Tabla 1. Descripción general de DAAs contra proteínas blanco del HCV aprobados por la FDA.	20
Tabla 2. Secuencias de Oligonucleótidos utilizados para amplificar las regiones que codifican la molécula de la inmunidad innata TICAM1 (5´ - 3´).	35
Tabla 3. Secuencia de Oligonucleótidos utilizados para amplificar las regiones que codifican la molécula de la inmunidad innata NOD1 (5´ - 3´).	36
Tabla 4. Descripción de la distancia media nucleotídica de la región HVR1 en pacientes con HCV en muestra basal.	47
Tabla 5. Descripción de los pacientes para el análisis filogenético de la región HVR1 a lo largo del tratamiento.	48

ABREVIATURAS

5B-SL3	Bucle 3´cis activo en la replicación de la proteína NS5B.
BOC	Boceprevir.
cDNA	DNA complementario.
emPCR	PCR en emulsión.
EVR	Respuesta virológica rápida (<i>Early virologic response</i>).
ER	Respuesta temprana (<i>Early response</i>).
ETR	Respuesta al finalizar el tratamiento (<i>End of treatment response</i>).
DAAs	Agentes antivirales de acción directa (<i>Direct - acting antiviral agents</i>).
DVR	Respuesta virológica tardía (<i>Delayed Virological Response</i>)
FDA	<i>Federal Drug Administration</i>
HAV	Virus de la Hepatitis A (<i>Hepatitis A Virus</i>).
HBV	Virus de la Hepatitis B (<i>Hepatitis B Virus</i>).
HCV	Virus de la Hepatitis C (<i>Hepatitis C Virus</i>).
HDV	Virus de la Hepatitis D (<i>Hepatitis D Virus</i>).
HEV	Virus de la Hepatitis E (<i>Hepatitis E Virus</i>).
HVR	Región Hipervariable (<i>Hypervariable región</i>).
IFN	Interferón.

IRF	Factor Regulador del Interferón (<i>Interferon regulatory factor</i>).
ISDR	Región determinante a la sensibilidad del Interferón (<i>Interferon sensitivity determining region</i>).
ISG	Genes estimulados por IFN (<i>IFN- I stimulated genes</i>)
IRES	Sitio interno de entrada al ribosoma (<i>Internal Ribosome Entry Site</i>).
IBT	Instituto de Biotecnología, UNAM
IC	Intervalo de confianza
GAP	Programa GAP (<i>Genome Assembly Program</i>).
LR	Respuesta tardía (Late response)
Log	Logaritmo
MDA5	<i>Melanoma Differentiation Associated protein 5</i> .
MID	Identificadores Múltiples (<i>Multiplex Identifiers</i>).
NCBI	<i>National Center for Biotechnology Information</i> .
NIH	Institutos Nacionales de Salud, USA (<i>National Institutes of Health</i>)
NGS	Secuenciación de nueva generación (<i>Next Generation Sequencing</i>).
NK	natural killer cell
NOD1	<i>Nucleotide-binding oligomerization domain containing 1</i> .
ORF	Marco de lectura abierto (<i>Open Reading Frame</i>).
PRR	Receptores de reconocimiento de patrones (<i>Pattern Recognition Receptor</i>)
RE	Retículo endoplásmico.
RAV	Variantes virales asociadas a resistencia. (<i>Resistance Associated Variants</i>)
RIG1	<i>Retinoic Acid Inducible gene 1</i>
RNA	Ácido ribonucleico (<i>Ribonucleic acid</i>).

RVR	Respuesta virológica rápida (<i>Rapid Virologic Response</i>).
SFF	Fluorogramas (<i>Standard Flowgram Format</i>).
SVR	Respuesta virológica sostenida (<i>Sustained Virological Response</i>).
SNP	Polimorfismo de nucleótido único (<i>Single Nucleotide Polymorphisms</i>).
SOF	Sofosbuvir.
TICAM 1	<i>Toll-like receptor adaptor molecule 1</i> .
TLR	Receptor tipo Toll (<i>Toll-like Receptor</i>)
UI/mL	UI/mL (“Unidades Internacionales por mililitro de sangre”).
UTR	Región no traducible (<i>Untranslated Region</i>).

RESUMEN

La Hepatitis C es una enfermedad causada por la infección con el virus que lleva el mismo nombre (HCV). En la respuesta inmune innata ante HCV participan diferentes vías de señalización: se ha descrito que diversas moléculas como TICAM1 (Toll-like receptor adaptor molecule 1) y NOD1 (Nucleotide-binding oligomerization domain containing 1) contribuyen a la correcta inducción y producción de interferones (IFN) tipo I, que son de crucial importancia en la respuesta antiviral. La identificación de Polimorfismos de Nucleótido Único (SNPs) presentes en regiones codificantes de estas moléculas de la inmunidad innata, permiten ubicar posibles marcadores moleculares de asociación en la infección por HCV, ya que variaciones en estos SNPs, pueden dar lugar a cambios de aminoácidos y llevar a la disminución o pérdida de la función de la proteína, lo que a su vez tiene repercusiones en la activación de los factores reguladores del interferón y en su inducción, afectando con ello la respuesta inmunológica contra HCV.

Aunado a lo anterior, el virus posee diversos mecanismos de variación genética que dan lugar a la generación de una amplia gama de subpoblaciones virales (cuasiespecies); en este sentido, se ha descrito que en la región HVR1 localizada entre las regiones que codifican las proteínas estructurales E1 y E2, se presenta la mayor variabilidad genética viral.

El objetivo de este proyecto fue caracterizar polimorfismos de nucleótido único en moléculas de la inmunidad innata implicados en la generación de cambios sinónimos y no sinónimos y establecer si tienen asociación con el tipo de respuesta al tratamiento y con la diversidad genética viral, en pacientes con Hepatitis C sometidos a terapia triple con Interferon, Ribavirina y Boceprevir, para lo cual se diseñaron una serie de oligonucleótidos que cubren regiones codificantes específicas en TICAM1 y NOD1 los cuales se secuenciaron por Sanger; simultáneamente, se realizó la caracterización nucleotídica de la región HVR1 en 4 pacientes mediante ultra secuenciación en la plataforma 454 de Roche.

El análisis permitió la identificación de dos SNPs que presentaron cambios sinónimos en la región codificante de la molécula TICAM1: en el rs2292151(C→T) se identificó el genotipo CC y en el SNP rs7255265 (G →A) el genotipo AA en el paciente con la respuesta más eficiente al tratamiento; contrario a los otros 6 pacientes que presentaron el genotipo CT (rs2292151) y GG (rs7255265); para la molécula NOD1 se identificaron dos SNPs implicados en cambios no sinónimos: en el rs2075820 (G→A) en un paciente con respuesta a las 12 semanas se identificó el genotipo heterocigoto GA, en el que se predice un cambio de ácido glutámico a lisina y en el rs376876391(G→A) se identificó el genotipo heterocigoto GA en el que se predice un cambio de arginina a histidina; los demás pacientes presentaron el genotipo homocigoto GG en ambos SNPs (rs2075820 y rs376876391).

Se realizó un análisis estadístico para evaluar el impacto de estos SNPs mediante un modelo de correlación con la prueba Tau B Kendall, encontrando una correlación moderada entre la severidad de la hepatitis (0.50) y la respuesta al tratamiento (0.65) en los SNPs presentes en la molécula TICAM1 y una correlación baja negativa entre los SNPs en la molécula NOD1 y la respuesta favorable al tratamiento (-0.39).

En cuanto al análisis de la región HVR1, el cálculo del promedio de las distancias nucleotídicas (\bar{d}) de las 4 muestras basales demostró que el paciente que inicialmente respondió y al final del tratamiento tuvo una pérdida de la respuesta (ID2), presenta la mayor

diversidad basal (22.55), en tanto que el paciente con la respuesta de eliminación viral más eficiente (ID7) es el que presenta la menor diversidad (3.68); aunado a lo anterior, el paciente no respondedor que presenta respuesta parcial, tiene una diversidad nucleotídica elevada (20.09), pero menor que el ID2, y el paciente respondedor a las 12 semanas (ID3) presenta una diversidad menor que los no respondedores (16.87). Lo anterior sugiere que la existencia de heterogeneidad genética entre los pacientes antes de iniciar el tratamiento, medida como π pudiera tener utilidad como marcador biológico de predicción al éxito del tratamiento antiviral.

Este trabajo describe por primera vez SNPs en moléculas de la inmunidad innata que podrían ser utilizado como moléculas asociadas a la respuesta ante la infección por HCV, además de la identificación de diferencias en la diversidad nucleotídica basal de las poblaciones virales de HCV antes de iniciar el tratamiento, lo que sugiere su potencial uso como biomarcador molecular para la generación de nuevos modelos de predicción a las terapias actuales contra HCV.

ABSTRACT

Hepatitis C is a disease caused by infection with the virus that bears the same name (HCV). In the innate immune response to HCV different signaling pathways are involved; It has been described that several molecules of the innate immunity as Toll-like receptor adapter molecule 1 (TICAM1) and Nucleotide-binding oligomerization domain containing 1 (NOD1) contribute to the successful induction of type I interferon (IFN) production, which is crucial in the antiviral response. Identification of single nucleotide polymorphisms (SNPs) at the codifying regions of these molecules, can provide insight of new molecular markers for association with HCV infection outcome, since changes at SNP level can lead to amino acid changes and to the reduction or loss of protein functionality, which can impact the activation of interferon regulatory factors and their induction, thereby affecting the immune response against HCV.

Furthermore, HCV virus has various mechanisms of genetic variation that give rise to the generation of a broad range of viral subpopulations (quasispecies), in this sense, the largest HCV genetic variability has been described in the HVR1 region, located between the E1 and E2 structural protein coding regions.

The aim of this project was to characterize SNPs in molecules of the innate immunity involved in the generation of synonymous and non-synonymous changes and establish if they have any association with the type of response to treatment and viral genetic diversity in patients with Hepatitis C undergoing triple therapy with Interferon, Ribavirin and Boceprevir; for this purpose, a series of oligonucleotides covering specific coding regions in TICAM1 and NOD1 were designed and sequenced by Sanger; Simultaneously, the nucleotide characterization of the HVR1 region was performed in 4 patients by ultra-sequencing on the 454 Roche platform.

Analysis allowed to the identification of two SNPs that showed synonymous changes in the coding region of the TICAM1 molecule: in the rs2292151 (C→T) was identified genotype CC and in the SNP rs7255265 (G → A) the AA genotype in the patient with the most efficient treatment response; contrary to the other 6 patients who presented the CT genotype (rs2292151) and GG (rs7255265); for the molecule NOD1, two SNPs involved in non-synonymous changes were identified: for the rs2075820 (G→A), the GA genotype in a patient with response at 12 weeks, in which a change from glutamic acid to lysine and for the rs376876391 (G→A) the heterozygous genotype GA was identified, in which is predicted a change of arginine to histidine; The other patients presented the GG homozygous genotype in both SNPs (rs2075820 and rs376876391).

Statistical analysis was performed to assess the impact of these SNPs using a correlation model with B Kendall Tau test. A moderate association was found between severity of hepatitis (0.50) and the response to treatment (0.65) in the TICAM1 SNPs, a low negative correlation at the NOD1 SNPs and favorable response to treatment (-0.39).

In regard of HVR1 region analysis, the calculation of the mean nucleotide distances (π) of the 4 baseline samples showed that, the patient initially responder but with loss of the response by the end of treatment (ID2), presented the highest basal diversity (22.55), while the patient with the most efficient viral elimination response (ID7) is the one with the lowest diversity (3.68), in addition to the above, the non-responding patient with partial response has a high nucleotide diversity (20.09), but less than ID2, and the responding patient at 12 weeks (ID3) has a lower diversity than non-responders (16.87). This suggests that the

existence of genetic heterogeneity among patients before starting treatment, measured as π , could have utility as a predictive biological marker for successful antiviral treatment.

This work describes for the first-time SNPs in molecules of innate immunity that could be used as molecules associated with the response to HCV infection, in addition to identifying differences in baseline nucleotide diversity of HCV viral populations before starting treatment, suggesting their potential use as a molecular biomarker for the generation of new prediction models for current HCV therapies.

1. Introducción.

1.1 Hepatitis C.

El término hepatitis, proviene del griego *hepar*, que significa hígado, fue utilizado por primera vez en 1710 y se refiere a todas aquellas enfermedades que producen una inflamación en el hígado (Du y Tang, 2016). La causa más frecuente que provoca hepatitis es una infección viral, aunque también puede ser producida por agentes químicos, el uso de fármacos, anestésicos, drogas, por una infección bacteriana o parasitaria, entre otras; por lo anterior, las hepatitis se dividen en infecciosas y no infecciosas. La mayoría de los pacientes refiere hepatitis de tipo infecciosa (Gómez I, 2003).

Las hepatitis virales son enfermedades infecciosas frecuentes, que representan un grave problema de salud pública; en el ser humano es una patología que puede ser caracterizada por inflamación y necrosis hepatocelular y pueden ser causadas por distintos virus hepatotrópos, la cual se presenta de forma aguda o crónica, dependiendo del agente infectante. Los agentes etiológicos causantes de hepatitis en el humano son los virus de la hepatitis A (HAV), virus de hepatitis B (HBV), virus de hepatitis C (HCV), virus de hepatitis D (HDV) y virus de la Hepatitis E (HEV) (Cordeiro, Taroco y Chiparelli, 2008); la información epidemiológica con respecto a las infecciones por HDV y HEV en la República Mexicana son escasas y poco frecuentes y los estudios que señalan su prevalencia sugieren una asociación con programas de higiene deficientes (Panduro et al., 2011). En la década de 1940, ya habían sido descritas dos tipos de hepatitis, la A y la B; en 1974 se reportó que cerca del 25% de los pacientes sometidos a cirugía desarrollaron una hepatitis post-transfusional y el 18% de ellos dieron negativo a los marcadores de ambos virus, además de presentar características clínicas diferentes, razón por la cual se denominó hepatitis no-A no-B. Fue hasta 15 años después que George Kuo, Qui-Lim Choo y Michael Houghton lograron identificar el agente causal responsable de la hepatitis no-A no-B al cual nombraron como hepatitis C (Du y Tang, 2016; Gómez I, 2003).

La hepatitis C es una enfermedad contagiosa causada por la infección con el virus del mismo nombre; el cuadro clínico, de acuerdo con la Organización Mundial de la Salud (OMS), puede ir desde un padecimiento leve de curso agudo, hasta una afección crónica que conlleva el riesgo de sufrir cirrosis hepática (20% de los pacientes) o cáncer del hígado (10% de los pacientes) los cuales se presente entre 20 a 30 años después de haber adquirido la infección; lo anterior convierte a la hepatitis C en un serio problema de salud pública en la actualidad (Gómez I, 2003; Preciado et al., 2014; Wilkins, Malcolm, Raina y Schade, 2010).

1.2 Agente Etiológico.

El HCV pertenece al género *Hepacivirus* de la familia *Flaviviridae*, es un virus esférico, envuelto, para el que se han descrito 7 genotipos virales y más de 150 subgenotipos. Su partícula viral completa tiene un diámetro aproximado de 55 a 60 nm (Figura 1).

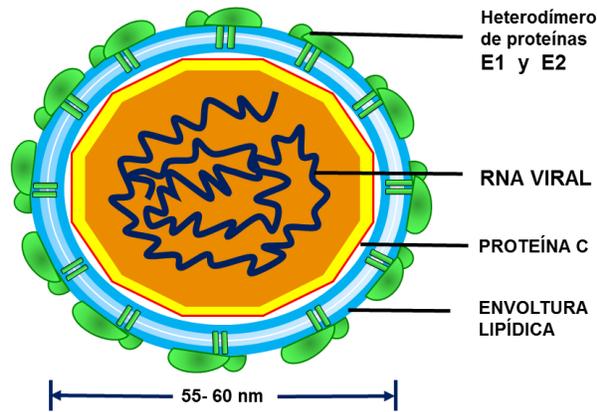


Figura 1. Estructura del HCV.

El genoma viral (Figura 2) está constituido por una cadena sencilla de RNA de polaridad positiva de aproximadamente 9,600 nucleótidos de longitud que codifica, mediante un único marco de lectura abierto (Open Reading Frame, ORF), para una poliproteína de 3,006 a 3,037 aminoácidos a partir de la cual se producen 10 proteínas funcionales: tres estructurales (C, E1 y E2), dos esenciales para la formación del virus (p7 y NS2) y cinco no estructurales (NS), que participan en los procesos de replicación del RNA viral (NS3, NS4A, NS4B, NS5A y NS5B) (Du y Tang, 2016; Lindenbach y Rice, 2013; Preciado et al., 2014).

Los extremos 5' y 3' son secuencias no codificantes (Untranslated Region, UTR) que flanquean la ORF. La región 5' UTR contiene 341 bases y es altamente conservada (Moradpour, Penin y Rice, 2007), contiene sitios importantes para la traducción, replicación y ensamblaje del genoma y su principal función es permitir la unión del ribosoma de la célula hospedera al RNA viral en la estructura conocida como IRES (Internal Ribosome Entry Site). El extremo 3' UTR es de 27-51 bases y participa de manera significativa en mejorar las señales de traducción e iniciación de la síntesis de RNA viral (Bartenschlager, Lohmann y Penin, 2013; Du y Tang, 2016).

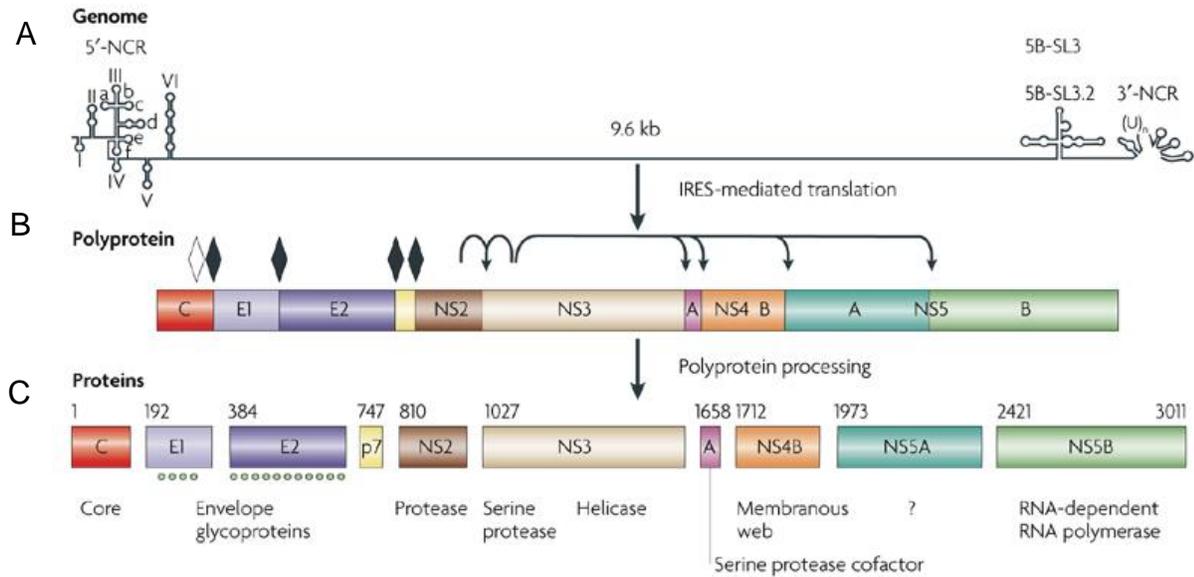


Figura 2. Organización genética y procesamiento de la poliproteína del HCV.

(A) Genoma de RNA de polaridad positiva de 9.6 kb, se indican las estructuras secundarias, las regiones no codificantes UTR 5' y 3', así como el bucle 3'cis activo en la replicación de la proteína NS5B (5B-SL3). (B) Procesamiento de la poliproteína que contiene el sitio interno de entrada al ribosoma (IRES) a partir del cual, se generan las proteínas estructurales y no estructurales; los diamantes de color negro indican los sitios de corte por la peptidasa en el RE y el rombo blanco indica el sitio de corte por la peptidasa señal de retículo endoplásmico en el extremo C- terminal de la proteína C. Las flechas curvas indican los sitios de corte por las proteasas virales NS2-3 y NS3-4A. (C) Representación de las 10 proteínas de HCV, los puntos en las proteínas E1 y E2 indican la glicosilación de las proteínas de la envoltura (4 y 11 glicanos N- ligados) (tomada de (Moradpour et al., 2007).

Algunas de las funciones de las proteínas del HCV ya han sido descritas, así, la proteína C constituye la núcleo-cápside viral, altamente conservada, que se une al RE, consta de 191 aminoácidos y 3 dominios con base a su hidrofobicidad, el primer dominio (comprende los aminoácidos 1 al 117) contiene residuos básicos, con dos regiones hidrofóbicas cortas y puede unirse al RNA viral; el dominio 2 (aminoácidos 118 al 174) es menos básico y más hidrófobo, por último, el dominio 3 (aminoácidos 175 al 191) es altamente hidrofóbico y actúa como una señal para la proteína de la envoltura E1 (Ashfaq, Javed, Rehman, Nawaz y Riazuddin, 2011; Du y Tang, 2016). La proteína C se encuentra involucrada en hepatocarcinogénesis y esteatosis de la hepatitis, ya que interacciona con numerosas proteínas celulares, afecta las funciones celulares tales como la transcripción de genes, el metabolismo de lípidos, apoptosis y varias vías de señalización (Du y Tang, 2016; Gómez I, 2003).

La glicoproteína de envoltura del virión E1 posee 192 aminoácidos. En el extremo carboxilo terminal (C-terminal) contiene de 4 a 5 sitios de glicosilación y es hidrofóbica, posee dos secuencias homólogas a segmentos transmembranales (Ashfaq et al., 2011; Gómez I, 2003).

La proteína E2 es otra glicoproteína de envoltura con 327 aminoácidos y contiene 11 sitios de glicosilación. Posee dos regiones hipervariables denominadas HVR1 y HVR2, que están sometidas a presión selectiva debida a la respuesta del sistema inmune ya que es la principal proteína blanco de anticuerpos neutralizantes. Numerosos estudios han puesto en manifiesto la heterogeneidad genética de la región HVR1, concluyendo que los cambios generados en esta región permiten al virus evadir el sistema inmune y facilitar la infección crónica (Ashfaq et al., 2011; Cortes et al., 2014; Cruz-Rivera et al., 2013).

Las proteínas E1 y E2 son procesadas por peptidasas del hospedero y forman un heterodímero que se expresa en la superficie de la membrana viral y sirve como proteína de unión a los diversos receptores celulares identificados hasta el momento, entre los que se encuentran: CD81, SRBI (Scavenger receptor class B type I), heparan sulfato y las lectinas de unión a manosa DC-SIGN (CD209) y L-SIGN (CD209-L) (Barth et al., 2003; Lozach et al., 2003; Scarselli et al., 2002).

El polipéptido P7 se encuentra codificado entre las proteínas estructurales y las no estructurales, es cortado de la poliproteína también por acción de peptidasas del hospedero, se desconoce su función pero en sistemas de membranas artificiales *in vitro* funciona promoviendo la translocación de NS2 en el RE para la escisión apropiada por peptidasas, estas proteínas forman canales iónicos que juegan un papel esencial en la infección viral teniendo papeles importantes en el ensamblaje y liberación de los viriones (Ashfaq et al., 2011; Griffin et al., 2004; Sakai et al., 2003).

El proceso proteolítico de la región no estructural es mediado por dos proteasas virales: la proteasa NS2 que rompe la unión NS2/3, y la proteasa NS3 dependiente de serina, que es la responsable de procesar los sitios NS3/NS4A, NS4A/B NS4B/5A y NS5A/5B (Figura 2). La proteína NS2 es un polipéptido transmembranal y su extremo carboxilo terminal se encuentra translocado en el lumen del RE, mientras el extremo amino se sitúa en el citoplasma. Su función no está totalmente definida, pero participa en el ensamblaje del virión y forma parte del complejo de replicación junto con la NS5A y NS5B (Ashfaq et al., 2011; Pallaoro et al., 2001).

La proteasa NS3 posee 631 aminoácidos, tiene un papel dual como serina proteasa (dominio amino-terminal) y como RNA helicasa (dominio carboxilo-terminal). La actividad de proteasa se incrementa por el cofactor NS4A que permite su completo plegamiento (Bartenschlager et al., 2013). El complejo NS3/4A es responsable del procesamiento de la poliproteína. La proteína NS3 resulta esencial para la generación de los componentes del complejo de replicación viral. Se ha descrito que el dominio serina proteasa de NS3 funciona como mecanismo de evasión de la respuesta inmune ya que produce la incisión en el aminoácido cisteína posición 372 de la proteína de la inmunidad innata TICAM1, que es crucial para la vía de

señalización de interferones tipo I (Z. Jiang, T. W. Mak, G. Sen y X. Li, 2004; Ovsyannikova et al., 2011). El conocimiento de diversos aspectos y mecanismos moleculares de esta enzima ha permitido el desarrollo de diversos fármacos que dan lugar a la inhibición específica de la misma (De Francesco y Migliaccio, 2005).

La proteína NS4 se divide en dos proteínas NS4A y NS4B; NS4A tiene 54 aminoácidos, en su extremo N- terminal contiene una región altamente hidrofóbica, mientras que el extremo C- terminal interactúa como cofactor con la proteína NS3 que forma parte del complejo de replicación viral. Por otra parte, NS4B es una proteína pequeña hidrofóbica, que participa en reclutar otras proteínas virales, contiene 4 dominios transmembranales y se ha reportado que una de sus principales funciones es inducir la formación de estructuras que soporten la replicación del RNA viral (Ashfaq et al., 2011; Einav, Elazar, Danieli y Glenn, 2004; Gómez I, 2003).

La región NS5 codifica para dos proteínas, la NS5A y NS5B. NS5A es una fosfoproteína hidrofílica que es de crucial importancia en la replicación viral, participa en la modulación de la señalización celular y la respuesta al interferón (IFN), ya que contiene una región implicada en la resistencia al IFN conocida como ISDR (Interferon sensitivity determining region), además, es capaz de regular la señalización celular e intervenir en mecanismos apoptóticos, por lo que se ha sugerido que tiene participación en la patogenia de la infección (L. Huang et al., 2005; Sarrazin, Herrmann, Bruch y Zeuzem, 2002). La proteína NS5B posee una longitud de 591 aminoácidos, actúa como RNA polimerasa dependiente de RNA y desempeña un papel importante en la síntesis de RNA viral. La estructura de la proteína permite el correcto plegamiento en 3 dominios, en uno de ellos se encuentra la actividad enzimática, en tanto los otros dos modulan la interacción con el RNA (Ashfaq et al., 2011; Bartenschlager et al., 2013).

1.3 Heterogeneidad y variabilidad genética del HCV

HCV posee 7 genotipos virales con una diversidad genética a nivel genotipo de aproximadamente un 30%. Los genotipos presentan una distribución característica que depende de la región geográfica; los genotipos 1, 2 y 3 poseen una distribución mundial, pero el genotipo 1 y 2 son endémicos de África occidental y el genotipo 3 de la India. Los genotipos 4 y 5 se distribuyen principalmente en África y el 6 en Asia, en tanto que la distribución del genotipo 7 no ha sido establecida (Preciado et al., 2014). Aunque existe una alta variabilidad genética viral, la diversidad de especies virales a nivel genotipo en diferentes regiones geográficas es relativamente pequeña, lo que sugiere una evolución a largo plazo (Preciado et al., 2014).

Además de los genotipos, existen que más de 150 subtipos, también llamados subgenotipos, que difieren aproximadamente en un 15% en su secuencia (Simmonds et al., 2005).

Los mecanismos por los que HCV genera su variación genética con diversos e incluyen mutación, deriva génica, recombinación e inserción de mutaciones puntuales por la RNA polimerasa dependiente de RNA. La evolución del virus a nivel intra hospedero da origen a una colección de diferentes variantes virales que se encuentran genéticamente relacionadas (denominadas cuasiespecies), las cuales evolucionan en un ambiente que permite la acumulación de sustituciones nucleotídicas con una alta tasa de frecuencia, calculada entre 1.4×10^{-3} y 1.9×10^{-3} sustituciones nucleotídicas por año, que representa una amplia distribución genética a lo largo del genoma, dando como resultado la heterogeneidad genética en HCV que contribuye a definir la enorme diversidad de la población viral, (Larrat et al., 2015; Sandres et al., 2000).

Esta capacidad del HCV de acumular sustituciones nucleotídicas es la que principalmente le permite escapar de la respuesta del sistema inmune y del tratamiento antiviral (Neumann et al., 1998; Preciado et al., 2014).

Otra mecanismo de evolución molecular en HCV son los cuellos de botella genéticos, que ocurren durante la transmisión y la fase aguda de la infección, en este mecanismo se reduce el tamaño de población viral en respuesta a la presión de selección ejercida por el sistema inmunológico, en consecuencia solo una fracción de variantes en la población viral logra reestablecer la infección en el hospedero (Bukh, Miller y Purcell, 1995; Preciado et al., 2014).

Mediante el uso tanto de técnicas convencionales como de nuevas metodologías de secuenciación, se ha logrado representar la variabilidad genética a lo largo de las regiones en el genoma que codifican para las proteínas del HCV (Figura 3), de esta forma se ha descrito que el gen que codifican para la proteína E2 presenta la mayor diversidad nucleotídica, es decir es considerado un gen heterogéneo, sobre todo en las regiones denominadas HVR (Hypervariable region) 1 y 2; la región HVR1 consta de una longitud aproximada de 27 aminoácidos y reporta una tasa de 0.87×10^{-3} sustituciones nucleotídicas por año; mientras que para la HVR2 (que consta de una longitud de 7 aminoácidos) hasta el momento no se tiene el reporte de la tasa de sustituciones nucleotídicas; datos recientes han sugerido que la heterogeneidad genética en las regiones HVR1 y HVR2 actúan en respuesta a diversas presiones de selección ejercidas en el hospedero, se ha descrito que las mutaciones en HVR1 permiten escapar de la presión ejercida por el sistema inmunológico (Farci et al., 1994; Hijikata, Kato, Ootsuyama, Nakagawa y Shimotohno, 1991; Weiner et al., 1992), estas mutaciones también se han asociado a la persistencia viral (Shimizu et al., 1994), a variaciones en el tropismo celular

(Qian, Camps, Maluenda, Civeira y Prieto, 1992), a una mayor progresión de la enfermedad hepática (Gretch et al., 1996; Naito et al., 1995) y recientemente se ha asociado a la respuesta al tratamiento (Cortes et al., 2014; Cruz-Rivera et al., 2013; Sandres et al., 2000). Es por lo anterior que la región HVR1 es representativa de las diferentes presiones selectivas a que es sometido el HCV.

Otra región que presenta alto grado de variabilidad es la región ISDR en NS5A, la cual ha sido asociada con diferentes grados de sensibilidad a la terapia convencional con interferón pegilado y ribavirina (IFN-RBV) dependiendo de las mutaciones virales presentes (Enomoto et al., 1996).

La región que se ha descrito con la menor variabilidad es la que codifica para la proteína C, sin embargo, un estudio ha demostrado que las sustituciones de aminoácidos presentes en esta región pueden ser utilizados como predictores de la respuesta al tratamiento con IFN-RBV (Akuta et al., 2007).

La región que codifica para la proteasa viral NS3, también presenta un grado relativamente alto de variabilidad genética. La comparación de secuencias de la región correspondiente al gen de la proteasa pertenecientes al mismo genotipo, pueden mostrar variaciones de hasta el 6% a nivel de la secuencia nucleotídica y hasta un 4% a nivel de aminoácidos (Larrat et al., 2015). En general, las secuencias consenso a nivel de aminoácidos de esta región tienden a ser más conservadas en aislados pertenecientes al genotipo 1, sin embargo, otros genotipos tales como el 3 muestran una mayor divergencia genética incurriendo en cambios fenotípicos que en ciertas ocasiones ocurren en el sitio activo de la enzima. Así, se puede inferir que el gen codificante para la proteasa viral tiende a estar altamente conservado entre aislados pertenecientes al mismo genotipo (Fusco y Chung, 2012).

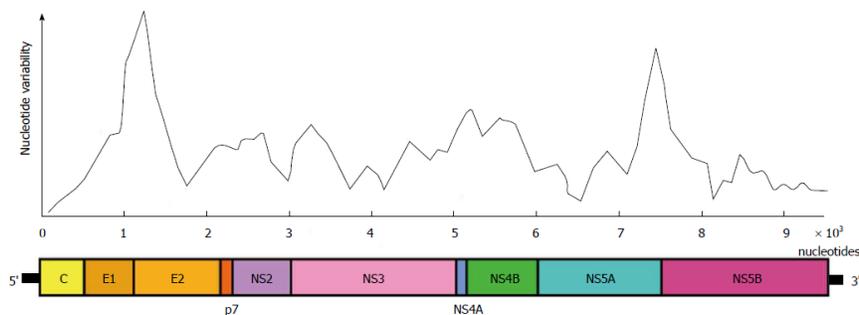


Figura 3. Genoma del virus y variabilidad nucleotídica del HCV.

Adaptado de (Preciado et al., 2014).

1.4 Transmisión del HCV.

La vía de transmisión de HCV es principalmente parenteral, mediante el contacto directo con los fluidos biológicos. La vía más frecuente y más demostrada es la transfusión sanguínea, seguido del uso de drogas intravenosas, de objetos punzocortantes contaminados, y en menor porcentaje se encuentra el trasplante de

órganos, la transmisión sexual, tatuajes y modificaciones corporales sin las prácticas de seguridad adecuadas, y la transmisión materno-fetal. Esta última depende del genotipo viral y carga viral de la madre infectada (Diago, 2004). El reconocimiento de las diversas vías de transmisión del HCV es de crucial importancia para la correcta implementación de mecanismos destinados a prevenir la propagación del virus (Gómez I, 2003; Preciado et al., 2014).

1.5 Progresión de la enfermedad por HCV.

La infección por HCV es infrecuentemente diagnosticada en la fase aguda. Las manifestaciones clínicas ocurren usualmente de 7 a 8 semanas después de la infección y hasta 2 a 26 semanas más tarde; no obstante, la mayoría de las personas permanecen asintomáticas, presentando solamente en algunos casos sintomatología mínima o leve cuyo principal signo clínico es la ictericia acompañada de dolor abdominal, fatiga, fiebre y náuseas. Se ha descrito la presencia de hepatitis fulminante en este periodo, pero es poco frecuente; en la mayoría de los casos (74 a 86%, aproximadamente) la infección se torna crónica y solo es diagnosticada cuando se presenta un cuadro clínico de cirrosis cuyo desarrollo puede ser de hasta 30 años (Franciscus, 2009).

La enfermedad crónica resulta de una infección aguda en la que no se logra eliminar al virus. En la mayoría de los casos, la inflamación crónica del hígado conlleva a diversos grados de fibrosis hepática, la cual puede estar acompañada sólo de fatiga o bien a complicaciones posteriores más severas como la cirrosis, que se desarrolla en un 15 a 20% de estos pacientes, e incluso la muerte. El tiempo de cada una de las fases de la enfermedad es variable, pero se calcula que un tercio de los pacientes puede desarrollar cirrosis de 20 a 30 años después de haber adquirido la infección (Díaz de Leon, 2003).

La evolución molecular de HCV, depende de diversos mecanismos de variación genética que se encuentran en el virus y las distintas presiones de selección, como las ejercidas por el sistema inmunológico dentro del hospedero, la vía de transmisión, zona endémica, tratamiento, edad, etc., que en conjunto, generan la gran diversidad de población viral, que pueden dar paso a formar diversas mutaciones distribuidas a lo largo de todo el genoma del HCV en respuesta a las presiones de selección ejercidas, dichas mutaciones han sido asociados en diversos estudios con la progresión de la enfermedad, ocasionando en la mayoría de los pacientes cuadros crónicos de fibrosis, cirrosis y hepatocarcinoma (Ripoli y Paziienza, 2011); sin embargo, los genotipos virales han sido asociados con la patogénesis de la infección, siendo el genotipo 1 el que mayormente se relaciona con la progresión de la enfermedad, aumento en el riesgo de fibrosis y carcinoma, resistencia a la insulina y mala respuesta al tratamiento; mientras que el genotipo 3 se asocia con el aumento de la esteatosis y fibrosis; es por ello que la identificación y caracterización de los genotipos y subtipos virales es un primer biomarcador de

soporte para inferir la progresión de la infección por HCV y la capacidad de respuesta al tratamiento (Preciado et al., 2014).

1.6 Epidemiología.

En el plano mundial, la OMS considera que existen alrededor de 185 millones de individuos infectados con Hepatitis C y en México, de acuerdo con los datos de los bancos de sangre, se estima una prevalencia de 1.5%, es decir hasta 2 millones de personas infectadas (I.C. 95% 1.3 – 1.7) (Burguete y cols., 2011; Kershenovich y cols., 2011) ubicándolo en un nivel epidemiológico intermedio. Las edades de mayor incidencia en la República Mexicana oscilan de los 41 a 50 años siendo el género masculino el que presentan mayor porcentaje de infección asociado con transfusiones sanguíneas y el uso de drogas intravenosas ($p < 0.001$) (I Burguete-García, J Conde-González y Jiménez-Méndez, 2011; Vera de León, 2005). (Ahmed y Felmler, 2015). El genotipo 1 viral (subgenotipo 1A, 33%; subgenotipo 1B, 21.4%) es el más frecuente en México, este tipo de virus tiene mayor resistencia al tratamiento por lo que es de vital importancia realizar estudios epidemiológicos constantes con el fin de identificar las frecuencias de este genotipo viral en el país. Se ha reportado en investigaciones recientes que la principal forma de transmisión por HCV en la población mexicana es la transfusión sanguínea con un porcentaje de 56% seguido del uso de tatuajes con un 18%; y el 16% asociado con drogas intravenosas (Espinoza et al., 2015). En la figura 4 se muestra la distribución de los genotipos en diversas regiones geográficas en la República Mexicana (I Burguete-García et al., 2011; Mendez, Salas, López, Cisneros y Castañeda, 2010; Sanchez-Avila, Gonzalez, Vazquez, Suarez y Uribe, 2007).

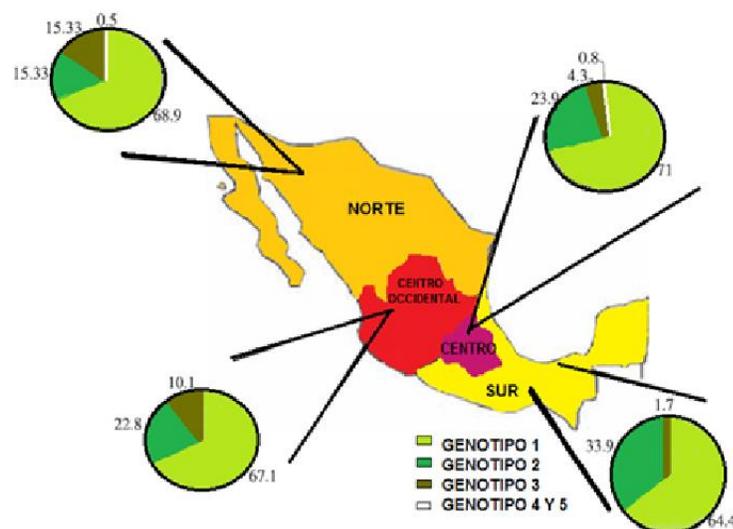


Figura 4. Distribución de los genotipos de HCV en la República Mexicana. Adaptado de (Sanchez-Avila et al., 2007).

1.7 Tratamiento.

Hasta el momento no existe una vacuna eficiente para prevenir la hepatitis C debido principalmente a la alta variabilidad genética del HCV (Bartenschlager et al., 2013).

Aunque la respuesta inmunológica es decisiva en la eliminación de agentes infecciosos, la mayoría de los agentes patógenos cuentan con mecanismos de evasión que restringen la eficiencia inmunológica. Esto puede dar lugar a la muerte del sujeto infectado, pero con mayor frecuencia al establecimiento de infecciones crónicas. Es aquí donde la inmunomodulación puede contribuir a resolver el problema al administrar fármacos que puedan estimular, complementar o restituir los elementos inductores o efectores de la respuesta que están siendo abatidos por la infección (Douam, Ding y Ploss, 2016). Los avances de la inmunología actual han provisto de conocimientos básicos que explican cada vez mejor los mecanismos y procesos inducidos por los patrones moleculares asociados a patógenos (PAMPs), en la respuesta innata, y los mediados por antígenos, en la respuesta adaptativa (Douam et al., 2016; Gokhale, Vazquez y Horner, 2014). Así, ya es posible combinar agentes terapéuticos con actividad antiviral o antimicrobiana con elementos de la respuesta inmunológica con muy buenos resultados.

En la actualidad, el tratamiento estándar en la República Mexicana para controlar la infección por HCV es la terapia combinada de Interferón Pegilado α (PEG-IFN) y Ribavirina (RBV).

La eliminación del virus por este tratamiento se conoce como Respuesta virológica Sostenida (Sustained Virological Response, SVR) que se define como la ausencia de RNA viral detectable en suero 24 semanas después de concluir el tratamiento (Hernández y Martínez, 2013). La SVR está estrechamente asociada con un riesgo disminuido para el desarrollo de cáncer hepático. No obstante, los estudios clínicos y epidemiológicos han señalado que la efectividad real de este tratamiento depende del genotipo viral infectante, con resultados del 76-82% en sujetos afectados por los genotipos 2 y 3, en tanto que con el genotipo 1 (que desafortunadamente es el más común en México) disminuye al 42-46% (Bosques-Padilla et al., 2003; Jimenez-Mendez y Castaneda-Hernandez, 2010).

La RBV, como análogo sintético de la guanosina, tiene propiedades antivirales a nivel de RNA, mediante la inhibición de la síntesis de RNA y consecuentemente de la replicación viral (Angelico et al., 2008). La ribavirina se fosforila intracelularmente para formar monofosfato (RMP), difosfato (RDP) y trifosfato (RTP), la incorporación de RTP por la RNA polimerasa lleva a la terminación e inhibición de la replicación, por lo que es considerado como un inhibidor de las polimerasas virales, al mismo tiempo que se incorpora a los IRES del genoma viral, inhibiendo la producción de la progenie de HCV (Durante-Mangoni, Iossa y Malgeri, 2011).

El PEG-IFN presenta la ventaja de tener una reducida excreción renal con respecto a la del IFN solo, lo cual incrementa su vida media en sangre (Angelico et al., 2008). Los interferones son proteínas naturales que estimulan el sistema inmunológico para combatir infecciones virales y tumores. Existen tres tipos de IFN producidos por el organismo en respuesta a una infección viral: los interferones tipo I que incluyen al IFN α y IFN β ; el interferón tipo II incluye solo al IFN γ ; y el interferón tipo III incluye al IFN λ con tres isoformas que implican a los genes IFN λ 1 (IL29), IFN λ 2 (IL28A) e IFN λ 3 (IL-28B) , los interferones tipo III están relacionados con los miembros de la IL-10 y la IL-22 que se ha demostrado que confieren hepatoprotección, además de inducir una cascada de transducción de señales muy parecida a la que inducen los IFNs tipo I, por tanto tienen actividades biológicas similares, ya que ambos tipos de IFNs tienen la habilidad intrínseca de inducir una respuesta antiviral (A. Balagopal, D. L. Thomas y C. L. Thio, 2010).

Existen dos presentaciones comerciales de IFN, PEG-IFN α 2a y PEG-IFN α 2b, que se prescriben dependiendo del tipo de genotipo viral. Generalmente en individuos con los genotipos 1 y 4, la terapia se establece con PEG-IFN α 2a y se les considera como grupo A. Por otro lado, aquellos pacientes con genotipo viral 2 y 3, son tratados con PEG-IFN α 2b y se establecen como grupo B (Graells, 2013).

Desafortunadamente este tratamiento tiene varias contraindicaciones y numerosos efectos secundarios, el grupo A llega a presentar depresión severa, prurito, hipertiroidismo y dermatitis; mientras que el grupo B cursa generalmente con fatiga, mialgia, depresión severa, náusea, alopecia, irritabilidad, hipertiroidismo, falta de apetito y dermatitis; además que en ambos grupos se presentan de forma frecuente anomalías en sangre como neutropenia, trombocitopenia y anemia (Bartenschlager et al., 2013).

Por lo tanto, se requieren terapias antivirales más eficaces, idealmente independientes de PEG -IFN α y Ribavirina. En la actualidad, están disponibles diferentes fármacos, denominados agentes antivirales de acción directa (DAAs) capaces de inhibir específicamente la función de ciertas proteínas virales de HCV indispensables para la replicación, entre ellos dos aprobados por la Federal Drug Administration (FDA) de los Estados Unidos de Norteamérica en 2011, el telaprevir y boceprevir, los cuales son inhibidores de la proteasa viral NS3/4a; así como el Sofosbuvir (SOF), inhibidor de la polimerasa viral, aprobado por FDA en 2013 en protocolo de terapia doble (en combinación con RBV para los genotipos 2 y 3) y terapia triple (SOF/PEG-IFN/RBV) para los genotipos 1 y 4; la incorporación de DAAs da lugar a tasas de SVR del 80% para los inhibidores de proteasa y mayor al 80% en los inhibidores de polimerasa (Thompson y McHutchison, 2009). Sin embargo, debido a que se ha descrito que la monoterapia con DAAs da lugar a la selección de variantes virales resistentes a los fármacos (Sarrazin y Zeuzem, 2010),

el uso del PEG-IFN y la RBV sigue siendo necesario en la terapia. En 2014, FDA aprobó una terapia dual libre de IFN a base de SOF y Ledipasvir (inhibidor de NS5a), así como el Viekira Pak (ombitasvir, paritaprevir y ritonavir + dasabuvir) para el tratamiento de pacientes infectados con el genotipo 1 aún con daño hepático avanzado (cirrosis compensada). Estos tratamientos libres de IFN, dan lugar a SVR superiores al 90%, sin embargo, se ha descrito la presencia de variantes virales naturales resistentes a estos fármacos que pueden emerger como las poblaciones virales dominantes en los casos donde no se logra una respuesta al tratamiento (Ahmed y Felmlee, 2015; Di Maio et al., 2014; Fonseca-Coronado et al., 2012).

FÁRMACOS APROBADOS POR LA FDA CONTRA HCV			
FÁRMACOS DAAS (Año Aprobado por la FDA)	PROTEÍNA BLANCO DEL HCV	COMPAÑÍA FARMACÉUTICA	MUTACIONES ASOCIADAS A RESISTENCIA A LOS FÁRMACOS EN LA PROTEÍNA BLANCO DEL HCV
TELAPREVIR (2011)	NS3/4 ^a	VERTEX	Q80K, V36A/M, T54A/S, R155K/T, A156 S/T, D168H/T/K/V/YN.
BOCEPREVIR (2011)	NS3/4 ^a	MERCK	
SIMEPREVIR (2013)	NS3/4 ^a	JANSSEN	
RITONAVIR (2014)	NS3/4 ^a	AddVie	
LEDIPASVIR (2014)	NS5A	GILEAD	M28A/G/T, L28T, F28S, L28F, Q30H/G/R/E/K, L31M/V, P32L/S, H58D, P58D, Y93C/H/N/S, Y93H.
OMBITASVIR (2014)	NS5A	AddVie	
PARITAPREVIR (2014)	NS5A	AddVie	
DACLATASVIR (2015)	NS5A	BMS	
SOFOSBUVIR (2013)	NS5B	GILEAD	S282T, V321I, L159F, L314H, C316H/Y, S368T, Y448C/H, A553T, A553V, G554S, S556R.
DASABUVIR (2014)	NS5B	AddVie	

Tabla 1. Descripción general de DAAs contra proteínas blanco del HCV aprobados por la FDA. (Ahmed y Felmlee, 2015; Lam, Jeffers, Younoszai, Fazel y Younossi, 2015; Pawlotsky, 2016).

1.8 Tipos de Respuesta al tratamiento.

La prueba diagnóstica de carga viral generalmente se realiza mediante el kit comercial COBAS® AmpliPrep/COBAS® Taqman® HCV Test, v2.0 (ROCHE) que permite cuantificar el número de partículas de RNA, expresado en UI/mL (“Unidades Internacionales por mililitro de sangre”) que se describe en logaritmos base 10 (log). Esta técnica se fundamenta en identificar al virus por la técnica de PCR en tiempo real mediante sondas de hidrólisis con oligonucleótidos específicos para la región

5´UTR de HCV (Zitzer et al., 2013). La prueba se emplea antes, durante y al finalizar el tratamiento, y la efectividad de la terapia se observará mediante la reducción del 99% de la carga viral o por el contrario el fracaso de la misma (Márquez, Santoscoy y Montoya, 2008; Mellor, Hawkins y Simmonds, 1999). La técnica posee un rango de detección para HCV de 15 IU/ml (1.18 log) a 100, 000, 000 IU/ml (8 log).

Existen diversos tipos de respuesta al tratamiento en asociación con la carga viral, que se conoce con el nombre de **terapia guiada por respuesta** ("EASL Clinical Practice Guidelines: Management of hepatitis C virus infection,"), clínicamente los resultados de la cinética viral son utilizados para continuar, extender o detener el tratamiento (Schijman et al., 2004), por lo anterior el criterio para finalizar la terapia es la **SVR** (Respuesta virológica sostenida), que se determina cuando la carga viral permanece no detectable 24 semanas después de finalizar el tratamiento, la SVR ha demostrado controlar la progresión de la infección por HCV en el 99% de los casos (Zitzer et al., 2013).

En esta clasificación se incluyen criterios como **RVR**, **EVR** y **DVR**, por lo que podemos decir que existe una **RVR** (Respuesta virológica rápida) cuando la carga viral no es detectable a las 4 semanas después de iniciar el tratamiento y hasta finalizar la terapia; se describe como **EVR** (Respuesta virológica temprana) cuando la carga viral disminuye a las 12 semanas de inicio de tratamiento y se mantiene por debajo de los valores detectables durante toda la terapia; en la **DVR** (Respuesta virológica tardía) los niveles de la carga viral disminuyen a la semana 12 de tratamiento, pero es hasta la semana 24 que son no detectables y permanecen hasta finalizar el tratamiento. Tanto la **RVR**, **EVR**, **DVR** generan una SVR que es clasificada como la respuesta exitosa al tratamiento (Figura 5) ("EASL Clinical Practice Guidelines: Management of hepatitis C virus infection,").

Sin embargo, la no respuesta al tratamiento se ha clasificado con los términos de **parcial**, **nula** y **perdida de la respuesta** (Figura 6); se dice que los pacientes tiene una respuesta **parcial** cuando existe una disminución de la carga viral mayor a 2 log a la semana 12 pero permanece detectable durante el tratamiento, respuesta **nula** es la caída de la carga viral menor a 2 log a la semana 12 tratamiento y continua siendo detectable hasta finalizar la terapia; por último la **perdida de la respuesta** hace referencia a un descenso de la carga viral, el virus permanece indetectable durante el tratamiento y al finalizar se detecta nuevamente ("EASL Clinical Practice Guidelines: Management of hepatitis C virus infection,").

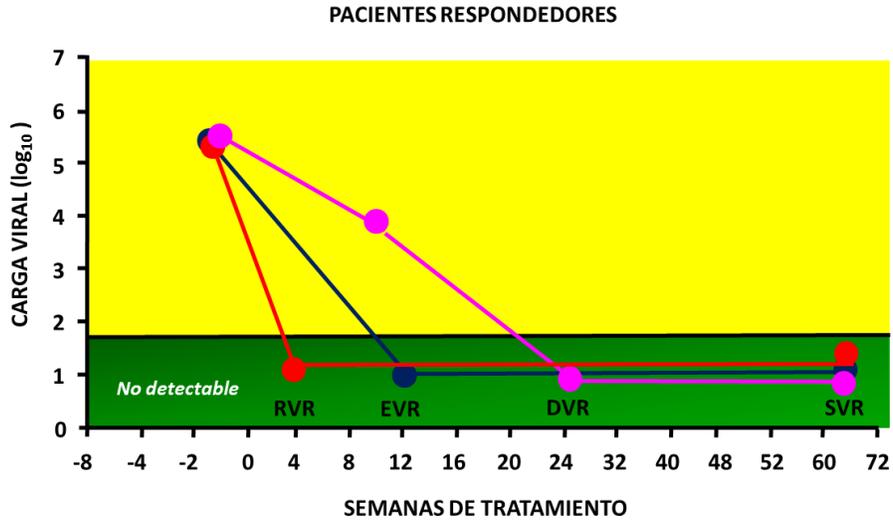


Figura 5. Gráfica de la carga viral en pacientes respondedores en HCV.

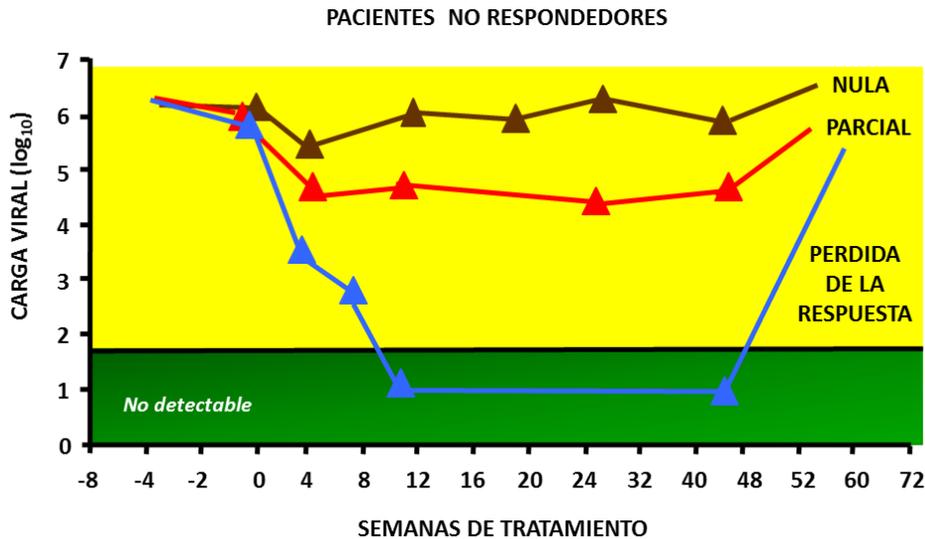


Figura 6. Gráfica de la carga viral en pacientes no respondedores en HCV.

1.9 Factores que intervienen en la respuesta al tratamiento.

Entre los otros factores que pueden influir en el tipo de respuesta al tratamiento pueden estar el genotipo del propio virus y las características genéticas relacionadas con el paciente.

Entre los factores relacionados con el virus, la existencia de los 7 genotipos virales y de más de 150 subgenotipos que generan la gran variabilidad genética viral, representan una limitante para un tratamiento exitoso; así, el genotipo 1 es el que menos responde al tratamiento (Preciado et al., 2014); por otro lado, la presencia de cargas virales elevadas (mayor a 5.77 log) son un factor de predicción igualmente negativo (Hernández y Martínez, 2013).

1.9.1 Polimorfismos de nucleótido único.

Hasta 2009, no se contaba con ningún indicador que pudiera predecir la respuesta al tratamiento antes de iniciar la terapia, situación que ha cambiado tras la identificación de la asociación de factores genéticos propios del paciente como son dos Polimorfismos de Nucleótido Único (SNPs) cercanos a la región promotora del gen del interferón lambda 3, también llamado IL-28B, han sido asociados con el éxito del tratamiento estándar la combinación de IFN y RBV (A. Balagopal et al., 2010; Tanaka et al., 2009; Thomas et al., 2009).

La asociación de los Polimorfismos de Nucleótido Único (SNPs) rs12979860 CC y rs8099917 TT cercanos a la región promotora del gen de IL28B junto con la SVR, constituyó una herramienta crucial como marcador molecular para predecir la respuesta al tratamiento dual (PEG-IFN-RBV) antes de iniciar la terapia; pero posee limitada predicción en los nuevos protocolos aprobados por la FDA con la incorporación de DAAs (Torres B., Venegas Santos y Brahm, 2011) y representa un antecedente muy importante que pone de manifiesto la necesidad de estudiar más ampliamente la estrecha relación entre el sistema inmune del hospedero, la capacidad de respuesta a los tratamientos y la generación de variantes virales resistentes a los mismos.

En este escenario, la identificación de SNPs presentes en moléculas de la respuesta inmune innata y la caracterización de las variantes virales del HCV asociadas a la resistencia a la terapia, puede contribuir de manera importante en la identificación de nuevos marcadores moleculares de respuesta al tratamiento y en el establecimiento de las interacciones entre las poblaciones virales y las moléculas del sistema inmune que determinan su eliminación o su permanencia en el hospedero.

Un polimorfismo de nucleótido único (SNP, Single Nucleotide Polymorphisms) se define como la variación que se presenta en una secuencia del DNA y que afecta sólo a una base nitrogenada (Figura 7). Para que estas variaciones sean consideradas como SNP, deben presentarse en un porcentaje mayor al 1 % de la población, de lo contrario, se considerara una mutación; es decir que un SNP posee una incidencia significativa en la población, por lo que tiene el potencial para funcionar como marcador biológico (Ramírez, Vargas, Tovilla y Frago, 2013).

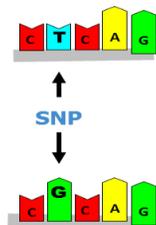


Figura 7. Esquema que representa un SNP.

Según su localización en el genoma, los SNPs se clasifican en: iSNP, si están localizados en regiones intrónicas; rSNP, si se presentan en regiones reguladoras; srSNP si afectan a la estructura y función de los RNAm (Ramírez et al., 2013); cSNP, aquellos presentes en regiones codificantes (exones); y en gSNP, localizados en regiones intergenómicas. Cada uno de ellos puede afectar, en último término, a la cantidad y actividad de las proteínas codificadas en los respectivos genes. Los cSNP pueden estar representados por SNPs sinónimos (sSNP) o no sinónimos (nsSNP) (Checa, 2007).

La identificación del genotipo viral, de los SNPs de IL-28, y de la dinámica de evolución intrahospedero de las especies virales, tanto en la terapia dual, como en pacientes con terapia triple, está generando nuevo conocimiento en la identificación de los mecanismos moleculares de la interacción virus-hospedero para su aplicación en la toma de decisiones médicas que contribuyan a mejorar la calidad de vida de los pacientes.

En este escenario, la identificación de SNPs presentes en moléculas de la respuesta inmune innata, puede contribuir de manera importante en la identificación de nuevos marcadores moleculares de respuesta al tratamiento y en el establecimiento de las interacciones entre las poblaciones virales y las moléculas del sistema inmune que determinan su eliminación o su permanencia en el hospedero.

En la respuesta inmune innata ante HCV participan diferentes vías de señalización, se ha descrito que diversas moléculas como Mitochondrial Antiviral Signaling protein (MAVS), Nucleotide-binding oligomerization domain containing 1 (NOD1) y Toll-like receptor adaptor molecule 1 (TICAM1), que intervienen e interaccionan en el reconocimiento de los componentes virales (Heim, 2013; Saeed, Mazoor, Jalal y Zahid Piracha, 2015; Zuniga, Macal, Lewis y Harker, 2015); la presencia de múltiples SNPs no sinónimos en estas moléculas, puede dar lugar a cambios de aminoácidos y llevar a la disminución o pérdida de la función de la proteína, lo que a su vez tendría repercusiones en la activación de los factores reguladores del interferón y en su inducción, con ello afectando la respuesta inmunológica contra HCV (Hise et al., 2015; Imran et al., 2012; K. Li et al., 2005). En este sentido se han descrito diversas asociaciones de polimorfismos presentes en moléculas de la inmunidad innata que tienen implicaciones importantes en la progresión a la infección por HCV, por ejemplo, el SNP R753Q que se encuentra en el TLR2 puede intervenir en el reconocimiento de la proteína core y NS3 de HCV y con ello mediar la insuficiencia hepática después de trasplante de hígado en el paciente (Brown et al., 2010); a su vez el gen KIR2DS3 de las células NK se asocia con el aumento de riesgo en la progresión crónica (OR 1.90, IC 1.25 – 2.90, $P < 0.002$), que en conjunto con el SNP rs12979860 genotipo TT (asociado a riesgo a la no respuesta al tratamiento de IFN y RBV (Tanaka et al., 2009; Thomas et al., 2009)) ubicado en el

gen de IL28B incrementa el riesgo crónico a la infección (OR 7.38, 95% IC 4.93-11.07, $P < 0.002$) (Dring et al., 2011).

1.9.2 Técnicas de identificación de SNPs.

Los métodos actuales para la determinación de SNP incluyen PCR en diversas modalidades y secuenciación de DNA. Los métodos genotípicos se han convertido en elementos esenciales en los análisis epidemiológicos, siendo poderosas herramientas que complementan las estrategias utilizadas para combatir la dispersión de las enfermedades infecciosas y una herramienta muy eficaz en el tratamiento (Vilchez y Alonso, 2009).

1.9.3 Secuenciación.

La secuenciación del DNA consiste en determinar el orden de las bases A, C, G y T en un fragmento de DNA. La secuenciación más utilizada hasta la fecha se realiza por el método descrito por Sanger en 1977 (Sanger, Nicklen y Coulson, 1977), permite obtener la secuencia de un fragmento determinado de DNA, un gen o parte de éste. Sin embargo, la alta demanda de secuenciación ha llevado al desarrollo de tecnologías de secuenciación masiva denominada como secuenciación de siguiente generación (Next Generation Sequencing, NGS) la cual tiene la capacidad de secuenciar en corto tiempo y a menor costo, genomas completos del hospedero y del microbioma presente en el organismo de un individuo. De esta manera se logra determinar el genomas tumorales o la existencia de polimorfismos, detectar genomas virales, como los del virus de la inmunodeficiencia humana y hepatitis (Jiménez, Gobernado y Sanchez, 2012).

El método enzimático de Sanger, utiliza una DNA polimerasa y terminadores que finalizan la cadena de DNA que está siendo sintetizada en lugares específicos, particularmente utiliza dideoxynucleótidos (es decir nucleótidos que en su carbono 3' no contiene el grupo hidroxilo). La incorporación de una base con esas características en una molécula nascente de DNA impide que una nueva base pueda incorporarse y la síntesis de DNA es interrumpida. Debido a que el ddT no contiene grupo 3' hidroxilo, la cadena no puede continuar extendiéndose, entonces la terminación ocurre específicamente en las posiciones donde dT debería haberse incorporado. Si un cebador y un molde son incubados con DNA polimerasa en presencia de una mezcla de ddTTP y dTTP, así como los otros tres deoxirribonucleosidos trifosfato (uno de ellos marcados radiactivamente o fluorocromos), se obtiene una mezcla de fragmentos todos con el mismo extremo 5' y con residuos ddT en el extremo 3'. Si esta mezcla es fraccionada por electroforesis desnaturizante en geles de acrilamida el patrón de bandas muestra la distribución de los residuos Timina en el DNA sintetizado. Utilizando terminadores con sus respectivos marcadores fluorescentes para cada uno de los nucleótidos, se determina la secuencia de bases (Ayllón, Cambra, Llave y Moriones, 2016;

Jiménez et al., 2012; Smith et al., 1986) La aparición de los secuenciadores automáticos permitió la instalación de centros de secuenciación, cuyo método general de trabajo (Figura 8) se realiza a partir de DNA molde, oligonucleótidos específicos para la región que se desea secuenciar más dideoxinucleótidos marcados con fluorocromos específicos para cada nucleótido en reacciones independientes por fluorocromo R6G (*N', N' - DIETIL- 2' 7' - dimetil- 6 carboxirrodamina*), R110 (*6-carboxirrodamina*), Tamra (*N, N, N', N' - tetrametil- 6 carboxirrodamina*), Rox (*6- carboxi- X rodamina*)); posterior a la desnaturalización de los fragmentos generados, los productos son detectados por electroforesis en un sistema computacional, cuyo resultado final es la obtención de electroferogramas almacenados en formatos con extensión .AB1, que son archivos de datos asociados con ABIF Applied Biosystems Inc. Format de la marca Applied Biosystems (Greif, 2006; Metzker, 2010).

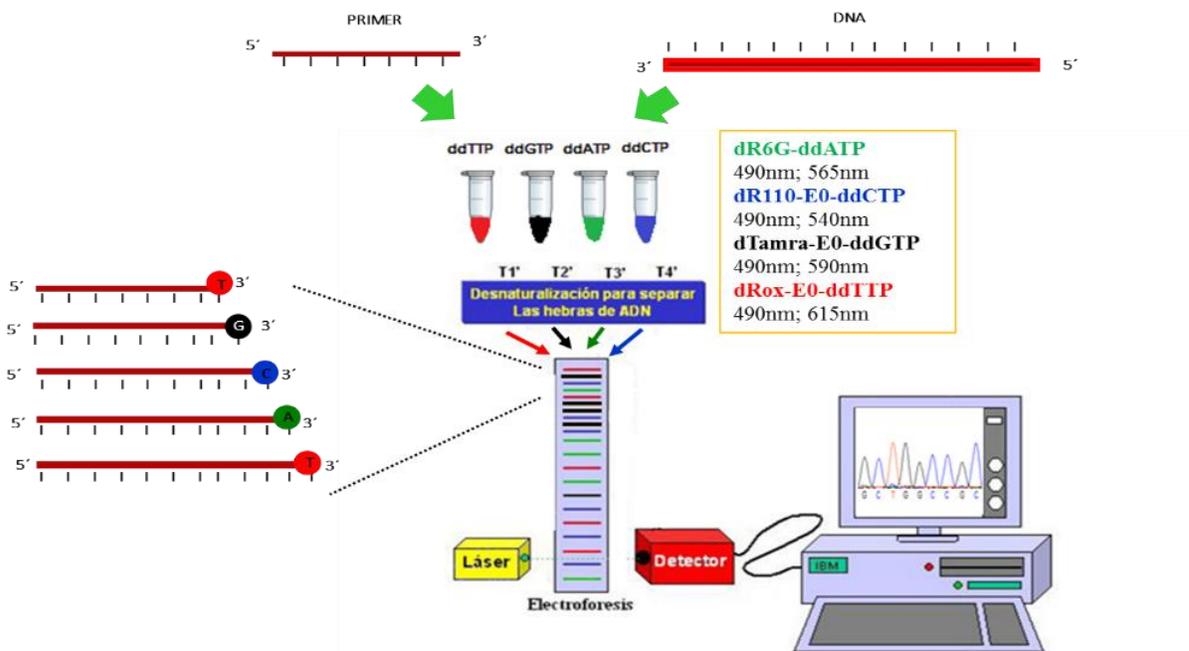


Figura 8. Diagrama general de trabajo en la técnica de secuenciación automatizada de Sanger.

1.9.4 Estrategias para el análisis bioinformático en la identificación de SNPs por secuenciación de Sanger.

Un electroferograma es el gráfico que almacena el secuenciador después de analizar la electroforesis, el electroferograma permite expresar el orden de las bases a partir de curvas de fluorescencia, una por cada base, y se puede visualizar en programas como sequence scanner (Lamerdin y Carrano, 1993), BioEdit (Hall, 1999) y el paquete Standen (Staden, Judge y Bonfield, 2003) que se obtienen de manera gratuita en internet. Es importante revisar la calidad de los

electroferogramas antes de continuar con los análisis de la muestra ya que es un parámetro que permite estimar la probabilidad de que cada uno de los nucleótidos secuenciados sean erróneos y genera confiabilidad de los resultados obtenidos (Rozo Murillo, Meneses Escobar y Franco Soto, 2011).

1.9.5 Paquete Bioinformático STADEN.

El paquete bioinformático Staden®, actualmente es de dominio público y proporciona una serie de herramientas para ensamblar, analizar y editar secuencias generadas por los secuenciadores automatizados. Para poder realizar el ensamble de secuencias, básicamente se puede utilizar dos programas Pregap y Gap, incluidos dentro del paquete (James K. Bonfield y Whitwham, 2010; Staden et al., 2003).

Pregap se utiliza para procesar los formatos nativos del secuenciador y sirve entre otras opciones para poder analizar la calidad de las secuencias a trabajar y arroja archivos de salida con formato SCF/ZTR, que permite continuar con el ensamble de secuencias. La versión original de *Gap* fue descrita en el año 1995 (J. K. Bonfield, Smith y Staden, 1995) es una interfaz que se utiliza fácilmente, el programa *Gap* (Genome Assembly Program) realiza un ensamble por secuencia en *contigs* largos (secuencias consenso que representa un conjunto de secuencias), la ventana principal del programa contiene el sector del contig, que se representa en una línea horizontal; sobre el contig se mostrará en conjunto la longitud del fragmento y el número de secuencias ensambladas; además de crear una base de datos que permite editar la secuencia seleccionando el contig. Una vez terminada la edición de la secuencia el programa exporta secuencias consenso de cada uno de los contigs largos generados y crear archivos en formato de texto extensión .FASTA (Andrews, 2008).

El diagrama general (Figura 9) del paquete Staden comienza incorporando los electroferogramas que arroja el secuenciador automatizado al programa PreGap que realiza un llamado de bases y filtra las secuencias en base a la calidad; posteriormente se realiza el ensamble de secuencias con el programa Gap4 que genera un alineamiento con secuencias consenso (contig), que permite editar secuencias y exportar archivos.

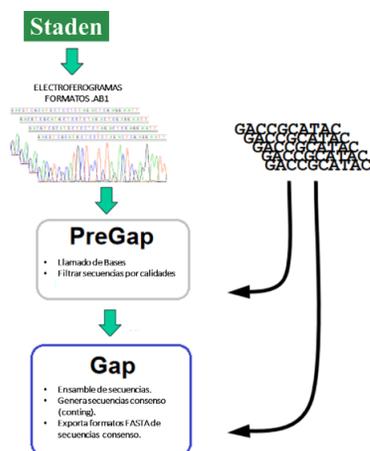


Figura 9. Diagrama general de trabajo del paquete Staden, Modificado de (Andrews, 2008).

1.9.6 Secuenciación Masiva en paralelo para identificar variantes virales intrahospedero de HCV.

La gran diversidad genética que presenta HCV; ha llevado a ocupar diversas tecnologías que permiten identificar a gran escala y profundidad variantes virales propias del hospedero, en este sentido la secuenciación masiva en paralelo permite reportar mutaciones puntuales a lo largo de toda la región del genoma que codifica para HCV (Park et al., 2014; Wang, Sherrill-Mix, Chang, Quince y Bushman, 2010).

Recientemente, la llamada secuenciación de segunda generación o NGS ha superado las limitaciones causadas por la secuenciación basada en el método de Sanger (Ayllón et al., 2016). La primera plataforma de secuenciación de nueva generación comercializada fue FLX Genome Sequencer de la empresa 454 Life Sciences en el año 2005, ahora propiedad del grupo F. Hoffmann - La Roche (Suiza). En 2007 Illumina Inc. (San Diego, California, USA) inició la venta del Genome Analyzer I (Bentley, 2006; Illumina, 2016) y Applied Biosystems (ABI) lo hizo con la plataforma SOLiD (AppliedBiosystems, 2008; Y.-F. Huang, Chen, Chiang, Chen y Chiu, 2012). Estas plataformas reducen de forma excepcional el tiempo y el costo económico, además que permiten revelar la secuencia de todo el material genético que contiene una muestra y facilita los estudios a gran escala de poblaciones y estudios epidemiológicos (Jiménez et al., 2012; Metzker, 2010).

La plataforma 454 (Roche) se basa en el uso de adaptadores específicos que se unen a los fragmentos de DNA que se van a secuenciar y se capturan en microesferas. Estas microesferas se introducen en una emulsión que contiene los reactivos necesarios para la PCR, se sumergen en aceite, generando un microambiente individual para cada una de ellas, posteriormente se lleva a cabo la amplificación en un termociclador; las microesferas con los productos amplificados se sitúan en una superficie de secuenciación (Metzker, 2005).

La tecnología 454 emplea una placa “picotiter” como soporte sólido y cada microesfera se coloca en un pocillo. Para iniciar la reacción de secuenciación, se añaden otras microesferas de látex que contienen las enzimas sulfúrilasa y luciferasa. En concreto esta técnica comienza con la obtención de una librería de pequeños fragmentos de DNA. En un primer paso se fragmenta el DNA a secuenciar mediante un proceso físico conocido como nebulización, donde el DNA se rompe en fragmentos de 200 a 800 pares de bases aproximadamente (Greif, 2006; Metzker, 2010).

La reacción de secuenciación para este caso, se realizan ciclos donde se ofrece en cada pocillo una base a la vez, secuencialmente. Durante la incorporación de una base en una molécula de DNA se libera pirofosfato, el pirofosfato liberado se convierte en luz mediante dos procesos enzimáticos (enzimas Sulfúrilasa y Luciferasa). La luz emitida en cada pocillo, donde se incorporó la base ofrecida, es monitoreada por la detección luminométrica de la liberación del pirofosfato durante la reacción de síntesis. Una cámara colecta los datos de cada base, en cada ciclo y en cada posición (Greif, 2006), tal como se visualiza en la figura 10, el flujo de trabajo que comienza con la aplicación de librerías del fragmento de DNA a secuenciar, continua con la generación de la librería e incorporación de adaptadores que permite realizar la amplificación clonar por PCR en emulsión, después cada microesfera contiene una copia de la librería que se coloca en una placa “picotiter” en el soporte sólido se realiza la corrida de secuenciación, para finalizar la reacción de secuenciación en el equipo genera los formatos nativos SFF que contienen la información nucleotídica secuenciada.

Ventajas:

- Muy rápido (hasta 700Mbps por ejecución)
- Más barato que Sanger
- Precisión bastante alta.

Desventajas:

- No puede generar secuencias muy largas (<1Kbs)
- Problemas para medir homopolímeros (secuencias de varios nucleótidos iguales).

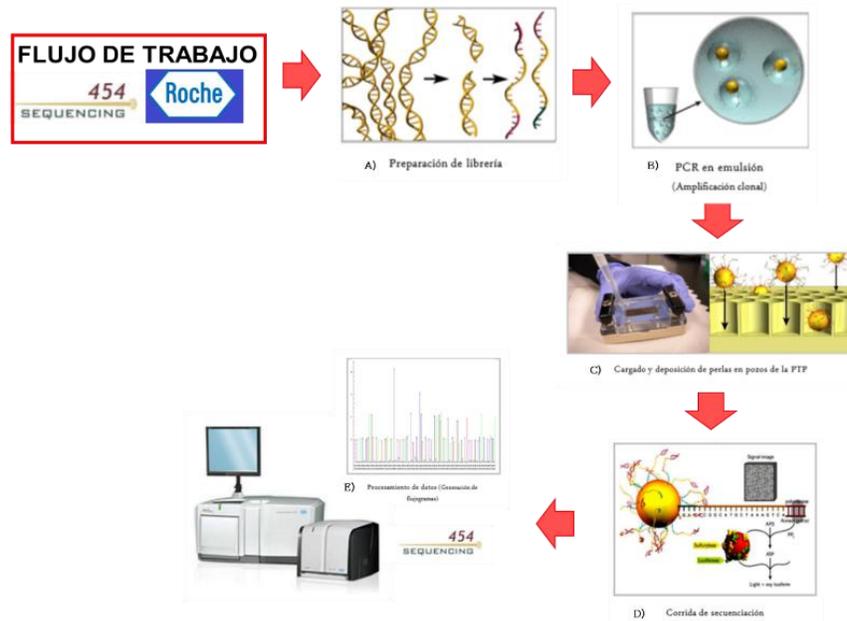


Figura 10. Diagrama general de trabajo de la plataforma 454 de ROCHE.

Por último, en la actualidad, se han desarrollado nuevas plataformas que constituyen la llamada tercera generación, se fundamenta en la amplificación de una molécula única de DNA, sin una amplificación clonal de las librerías previamente generadas; empresas como Pacific Biosciences (Carneiro et al., 2012) y Oxford Nanopore (Hargreaves y Mulley, 2015) ofrecen novedosas soluciones y rigen las nuevas plataformas.

1.9.7 Herramientas bioinformáticas para el análisis de datos por secuenciación masiva en paralelo.

La gran magnitud de los datos obtenidos con las plataformas NGS requiere el empleo de programas informáticos para evaluar los resultados y, como en muchas ocasiones, los avances tecnológicos han contribuido en el análisis para una apropiada interpretación de los resultados biológicos (Rozo Murillo et al., 2011); actualmente existe una amplia gama de algoritmos bioinformáticos y en general los softwares que rigen el análisis bioinformático pueden ser categorizados dentro de 3 etapas: primario, secundario y terciario (De Necoche y Canul, 2004).

El análisis bioinformático primario se define como los pasos específicos necesarios para que los secuenciadores generen los pares de bases y los puntajes de calidad expresados en archivos FASTQ (De Necoche y Canul, 2004); en tanto, el análisis secundario hace referencia al ensamble de las lecturas, incluyendo pasos previos de limpieza como el filtrado de calidad, el ensamble de secuencias que a su vez contiene dos tipos de análisis; la primera con referencia o mapping, en donde la secuencia del genoma tiene en cuenta la referencia del gen de interés; mientras la segunda se encarga de ensamblar lecturas sin disponer de un genoma de

referencia, lo que se denomina secuenciación de novo. La eficiencia de los algoritmos dependerá de la longitud de la lectura y la profundidad o cobertura de secuencia (Rozo Murillo et al., 2011).

Entre los programas más empleados para secuenciación de novo destacan:

- AbySS (Simpson et al., 2009).
- MIRA3 (Wilson et al., 2013).
- Velvet (D. R. Zerbino, 2010; Daniel R. Zerbino y Birney, 2008)

En el caso de programas para el análisis con disponibilidad de una secuencia de referencia destacan

- Bowtie (Langmead, Trapnell, Pop y Salzberg, 2009).
- BWA (H. Li y Durbin, 2009, 2010).
- MOSAIK (Lee et al., 2014).

No existe una herramienta bioinformática para cada tipo de análisis, por lo tanto, se deben conocer los puntos débiles y fuertes de cada uno de los programas, puesto que cada uno ha sido diseñado con un propósito específico. Cada plataforma ha desarrollado sus propias herramientas de análisis, por ejemplo, ROCHE ofrece a su vez el programa “GS De Novo Assembler, Reference mapper and Variant Analyzer” (Greif, 2006; Jiménez et al., 2012; Metzker, 2010).

Por último, el análisis terciario se encarga de darle sentido biológico a los datos obtenidos, como la agregación de datos, la anotación, análisis de estructura de los datos entre otros. Es la fase encargada de interpretar la información y obtener conclusiones (De Necoche y Canul, 2004).

1.9.2.1.1 Plataforma Galaxy.

Galaxy es una plataforma web (Giardine et al., 2005) que permite ejecutar herramientas bioinformáticas con la posibilidad de construir flujos de trabajo; combina la utilidad de las bases de datos de anotación de genomas existentes en un entorno amigable para buscar recursos externos (Afgan et al., 2016).

Las principales características de Galaxy son la accesibilidad, la reproducibilidad y la transparencia; con lo anterior permite a los usuarios acceder a los recursos de forma amigable sin la necesidad de ejecutar programación (accesibilidad), tiene la capacidad de almacenar la información sobre cada paso en el análisis computacional permitiendo repetir el experimento (reproducibilidad), finalmente Galaxy ofrece la opción de compartir públicamente o con particulares los objetos generados (transparencia) (Giardine et al., 2005); lo anterior hace que la plataforma sea una herramienta útil para el desarrollo de investigaciones en el área biológica.

2. Objetivos.

2.1 Objetivo General.

Caracterizar polimorfismos de nucleótido único en moléculas de la inmunidad innata implicados en la generación de cambios sinónimos y no sinónimos y establecer si tienen asociación con el tipo de respuesta al tratamiento y con la diversidad genética viral, en pacientes con Hepatitis C sometidos a terapia triple.

2.2 Objetivos Particulares.

- Determinar el genotipo de diversos SNP presentes en dos moléculas de la inmunidad innata (TICAM y NOD1) que representen cambios sinónimos y no sinónimos en pacientes con hepatitis C sometidos a terapia triple (Interferon pegilado, Ribavirina y Boceprevir) mediante secuenciación de Sanger.
- Establecer una relación entre los polimorfismos presentes en las moléculas de la inmunidad innata TICAM y NOD1 con la respuesta al tratamiento en pacientes infectados con Hepatitis C sometidos a terapia triple.
- Realizar la secuenciación de la región HVR1 en diferentes muestras de los pacientes infectados con hepatitis C sometidos a terapia triple mediante secuenciación masiva paralela en la plataforma 454 de Roche.

3. Justificación.

En la República Mexicana, el tratamiento estándar para controlar la infección por HCV es la terapia dual combinada de Interferon Pegilado (PEG-IFN) y Ribavirina (RBV), que tiene un costo aproximado de \$ 280,000.00 (MXN) por paciente (Younossi et al., 2016) y se aplica a todos los pacientes sin contraindicación que son atendidos en el sector público del sistema nacional de salud en México; sin embargo, dicha terapia combinada posee un porcentaje de respuesta en los pacientes de solo 45% de SVR; es decir, la mayoría de los pacientes no responde al tratamiento. En la actualidad, los agentes antivirales directos dan una tasa mayor a 80% de SVR, pero tienen un costo elevado (la terapia de PEG-IFN, RBV y BOC reporta un costo de \$1, 273, 800. 00 (MXN) por paciente; mientras el tratamiento combinado de PEG-IFN, RBV y SOF \$1, 805, 840. 00 (MXN) por paciente (Younossi et al., 2016).

La identificación del genotipo viral, de los SNPs de IL-28, y de la dinámica de evolución intrahospedero de las especies virales, tanto en la terapia dual, como en pacientes con terapia triple, genera nuevo conocimiento en la identificación de los mecanismos moleculares de la interacción virus hospedero para su aplicación en la toma de decisiones médicas que contribuyan a mejorar la calidad de vida de los pacientes. Sin embargo, las nuevas terapias libres de interferón, requieren la identificación de nuevos marcadores moleculares de respuesta al tratamiento. En este escenario, la identificación de SNPs presentes en moléculas de la respuesta inmune innata, puede contribuir de manera importante en la identificación de nuevos marcadores moleculares de respuesta al tratamiento y en el establecimiento de las

interacciones entre las poblaciones virales y las moléculas del sistema inmune que determinan su eliminación o su permanencia en el hospedero.

La identificación de variaciones en la secuencia de moléculas de la inmunidad innata que participan en la respuesta ante la infección por HCV no han sido descritas en la actualidad; los SNPs presentes en la secuencia codificante de estas moléculas, pueden alterar la función de las proteínas y, por lo tanto, modificar la respuesta inmunológica, generando una mayor capacidad para controlar la replicación viral o una menor capacidad de respuesta que permita la generación de variantes virales con el consecuente impacto sobre el tipo de respuesta al tratamiento.

4. Metodología.

La estrategia general de este trabajo consistió en realizar dos metodologías de secuenciación, la primera enfocada en caracterizar polimorfismos presentes en las moléculas TICAM1 y NOD1 en 6 pacientes con HCV sometidos a terapia triple (PEG IFN, RBV y BOC) por la técnica de Sanger (ID1, ID2, ID3, ID4, ID5 y ID7); el segundo método de secuenciación se llevó a cabo por NGS en la plataforma 454 de ROCHE para caracterizar la región HVR1 de 4 de los pacientes, dos respondedores (ID3 y ID7) y dos no respondedores al tratamiento (ID1, ID2) (Figura 11).

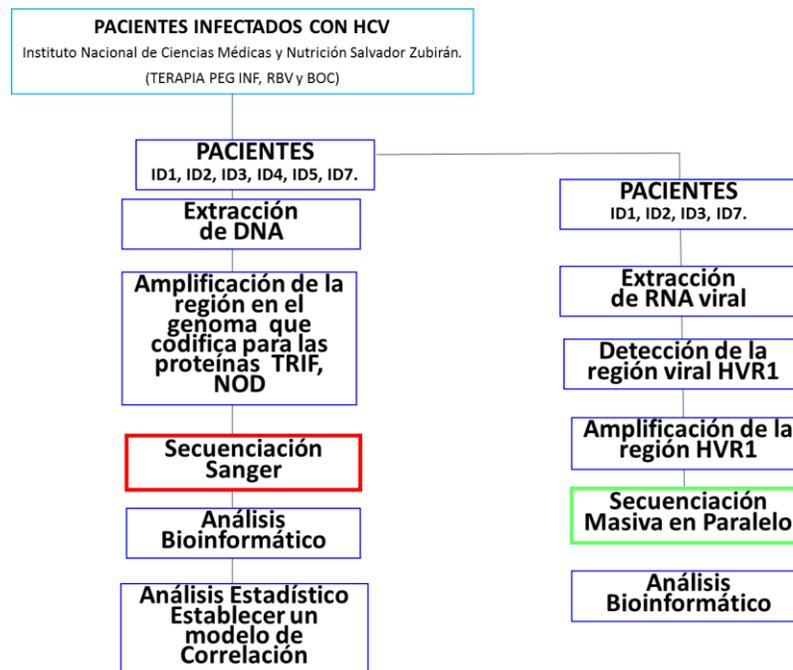


Figura 11. Estrategia general experimental de trabajo. Se observa la población de estudio 6 pacientes infectados con HCV sometidos a terapia triple.

Las metodologías realizadas: secuenciación de sanger (cuadro rojo) y secuenciación masiva en paralelo (cuadro verde).

4.1 Diseño de oligonucleótidos específicos para la identificación de las moléculas TICAM1 y NOD1.

Para lograr secuenciar y amplificar las regiones codificantes de las moléculas TICAM1 y NOD1 se realizó el diseño específico de oligonucleótidos en el programa PRIMER SELECT contenido en el paquete DNASTAR Lasergene. Se localizaron todos los SNPs que corresponden a cambios sinónimos y no sinónimos reportados en la base de datos del National Center for Biotechnology Information (NCBI), posteriormente se descargó la secuencia FASTA de referencia de la molécula TICAM1 del NCBI con número de acceso NM_182919.3 (Figura 12) y la NM_006092.2 (Figura 13) para la molécula de NOD1, que se visualizaron en el programa MEGA6.

Se diseñaron los oligonucleótidos cuyas principales características fueron una Tm de 60°C, una longitud de 20 – 25 bases nucleotídicas, para generar productos de 500 pares de bases y sobrelapados 100 pares de bases para evitar incidir sobre algún cambio de base. Los oligonucleótidos diseñados se observan en la tabla 2 para la molécula TICAM1 (6 pares de oligonucleótidos) y tabla 3 para NOD1 (3 pares de oligonucleótidos).

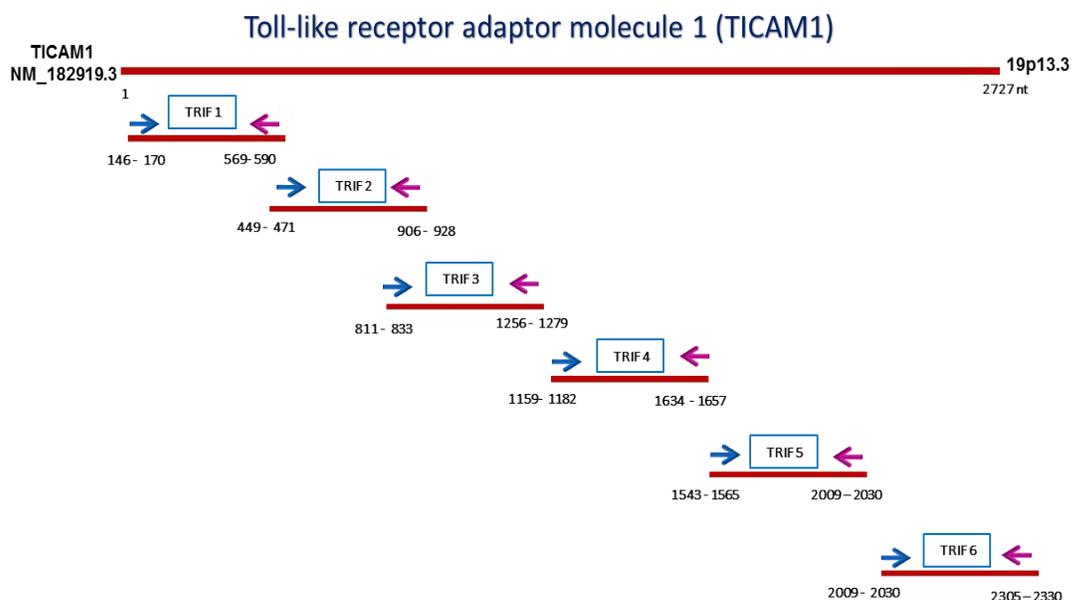


Figura 12. Estrategia del diseño de oligonucleótidos para la molécula TICAM1.

INICIADORES DISEÑADOS PARA AMPLIFICAR LA REGION QUE CODIFICA LA MOLÉCULA TICAM1.	
TICAM-F1 CCTGCAGAACCTGGAACAGTGAATG	TICAM-R1 GCTTCTCCTCAGCCAGCAGGTG
TICAM-F2 CTAGAGGCATTGAAGGCCGATGC	TICAM-R2 CAGGCTGAGGAAGGGCATGGTAG
TICAM-F3 CATTGACGGTGTTCGGACTGGA	TICAM-R3 AGAAAGTTGGAGTGGCGTCTGGTC
TICAM-F4 CACCAACTACCCAGTGGAGTGAC	TICAM-R4 GCGACAGTCGAAGTTGGAGGTGAG
TICAM – F5 CACCTTCTGCGAGGATTTCCAGG	TICAM – R5 GGAGCTGCTCCATCTGTGCCTG
TICAM – F6 CAGATCTTCGCCAGGAAGGTGG	TICAM – R6 GTACCATCTGTGCGTGGTGGATAATG

Tabla 2. Secuencias de Oligonucleótidos utilizados para amplificar las regiones que codifican la molécula de la inmunidad innata TICAM1 (5' - 3').

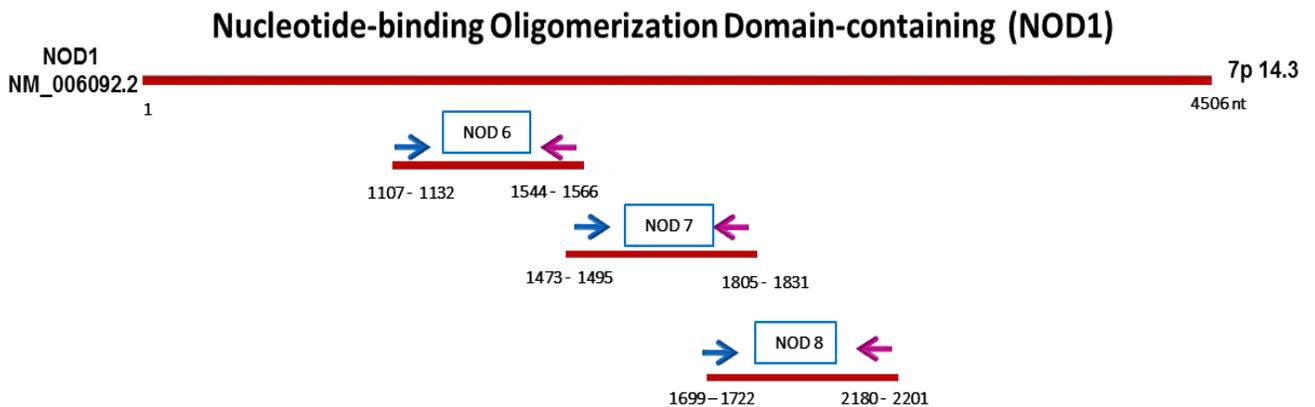


Figura 13. Estrategia del diseño de oligonucleótidos para la molécula NOD1.

INICIADORES DISEÑADOS PARA AMPLIFICAR LA REGION QUE CODIFICA LA MOLÉCULA NOD1.	
NOD-F3 GTGAGACCATCTTCATCCTGGGTGAT	NOD-R3 AGCACCTTCTCCGCAGGAACTG
NOD-F4 TGCTCAGTGGGAAGCTGCTCAAG	NOD-R4 CATCCTGTTTCAGATGGACCTCAGTGAC
NOD-F5 CTTCTGCTGGATCATCTCCGGTG	NOD-R5 GATAGCAGGACGTGGTCGCTGC

Tabla 3. Secuencia de Oligonucleótidos utilizados para amplificar las regiones que codifican la molécula de la inmunidad innata NOD1 (5´ - 3´).

4.2 Población de Estudio.

Se seleccionaron 6 pacientes (ID1, ID2, ID3, ID4, ID5, ID7) mexicanos con diagnóstico de infección de HCV confirmado por la presencia de anticuerpos anti-HCV y carga viral que fueron sometidos a terapia triple con PEG-INF, RBV y BOC bajo los protocolos del Departamento de Gastroenterología del Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán. Los criterios de inclusión fueron que el paciente aceptará participar en el estudio, que no tuviera otras enfermedades infecciosas o inmunodeficiencias y que contará con la siguiente información: tiempo de infección, historial clínico, monitoreo de carga viral, fecha de inicio del tratamiento antiviral, genotipo viral infectante.

4.3 Extracción de DNA y Genotipificación de SNPs en las regiones que codifican para las moléculas TICAM1 y NOD1.

El DNA de muestras de sangre total fue aislado mediante el empleo del estuche comercial QIAamp® DNA Blood Mini Kit (250) (Qiagen). El DNA purificado fue utilizado como molde para la amplificación de la región que codifica para las moléculas de la inmunidad innata TICAM1 y NOD1. La amplificación se llevó a cabo en un termociclador de tiempo real PCR thermocycler Stratagene Mx 3000P-02 por el método de agente intercalante (SYBR Green), los oligonucleótidos utilizados fueron diseñados como se muestra en la tabla 2.

La reacción de PCR se realizó bajo las siguientes condiciones: 5 min por 95°C, seguido por 40 ciclos de desnaturalización a 95°C por 10 s, alineación 60 °C por 10 s y extensión 72°C por 10 s, para incorporar una curva de disociación a 95°C por 1s, 60° C por 1 s y 95°C por 10 s.

Con el producto de PCR se corrió un gel en el sistema E-Gels size select ® para verificar el tamaño del producto y purificar el amplicon generado en cada caso.

Posteriormente se realizó la secuenciación por la plataforma de Sanger con el equipo comercial utilizado por el Instituto de Biotecnología (IBT, UNAM) incorporando primer forward y reverse por separado.

4.4 Análisis Bioinformático de la secuenciación de SNPs y predicción *in silico*.

Los formatos con extensión .AB1 que contienen los electroferogramas fueron analizados e incorporados en conjunto en una base de datos por el paquete Staden Package en el programa GAP4 para identificar, seleccionar y clasificar la lectura de secuenciación en expresión de bases nucleotídicas. Los formatos fasta generados en el paquete sirvieron como formatos de entrada para predecir la estructura *in silico* de la proteína en la aplicación Swiss-Prot (Biasini et al., 2014) y visualizados en Jmol (PyMOL, 2012).

4.5 Análisis Estadístico de la Secuenciación de SNPs.

Con los SNPs presentes en las moléculas TICAM1 y NOD1 se realizó una prueba de correlación Tau B de Kendall (criterio no paramétrico basado en intervalos jerarquizados), que permitió asignar un valor a los SNPs y visualizar la relación que existe entre la severidad de la infección por HCV, el genotipo de IL28B en los pacientes y la respuesta al tratamiento, generando una matriz que se ejecutó y graficó en el lenguaje de programación RStudio (RStudio, 2005) con los paquetes reshape2 y ggplot2.

4.4 Extracción de RNA Viral y Secuenciación de la Región HVR1 por NGS

La identificación de variantes virales (cuasiespecies) se llevó a cabo mediante un protocolo de secuenciación de segunda generación previamente desarrollado por el grupo de investigación de Fonseca-Coronado (Cruz-Rivera et al., 2013; Fonseca-Coronado et al., 2012). El RNA fue aislado a partir de suero de los pacientes (de los pacientes ID1, ID2, ID3 e ID7 se tomó una muestra basal; mientras que de los pacientes ID2 e ID3 las muestras seleccionadas fueron antes, durante y al finalizar el tratamiento) mediante el empleo del estuche comercial QIAamp® Viral RNA Mini Handbook (Qiagen). Con el material resultante se llevó a cabo la retrotranscripción con oligonucleótidos específicos para la secuencia que codifica la región HVR1 por el método de síntesis OneStep® RT-PCR kit (Qiagen); la reacción de transcripción reversa se llevó a cabo por 30 min a 50°C seguido por 15 min a 95°C, incorporando 40 ciclos de desnaturalización a 95°C por 5 s, alineación 50 °C por 5 s y extensión 72°C por 10 s y un ciclo de enfriamiento de 72°C por 1 min y 4°C por 10 min. La reacción de amplicon deep sequencing se realizó en un equipo 454 Genome Sequencer FLX de Roche para lo cual cada muestra fue amplificada independientemente con primers de fusión, incluyendo los primers adaptadores A y B del equipo 454 y un set de primers internos, acoplados a uno de los identificadores múltiples (MID), que son secuencias identificadoras únicas de diez nucleótidos que permiten marcar los amplicones asociados con cada una de las muestras para la

secuenciación múltiple. Los fragmentos generados de muestras individuales se amplificaron por separado y purificaron con AMPure beads (Agencourt), se cuantificaron por fluorimetría (Invitrogen; Quant-iT PicoGreen dsDNA) y se mezclaron a concentraciones equimolares. La amplificación clonal en perlas mediante PCR en emulsión (emPCR) se realizó utilizando los adaptadores del sistema 454 A y B que permitieron secuenciar ambas cadenas de DNA. La pirosecuenciación se llevó a cabo usando una placa PicoTiterPlate de 16 sectores. Los archivos conteniendo los fluorogramas (SFF) fueron procesados utilizando el programa GS Amplicon Variant Analyzer (Roche), se extrajeron los datos de secuencia de nucleótidos asociados con cada amplicon marcado con un MID diferente para subsecuentemente identificar las secuencias pertenecientes a cada muestra; la emPCR y la corrida de secuenciación fueron realizados por el Laboratorio de Servicios Genómicos Langebio, CINVESTAV, unidad Irapuato que entrego los formatos SFF para su posterior análisis.

4.6 Análisis Bioinformático de la región HVR1.

Para el análisis de los formatos nativos que arroja el secuenciador se utilizó el programa NextGen WorkBench (Workbench, 2015) un editor que convierte los formatos SFF a FastQ, por muestra analizada para limpiar y filtrar las secuencias en calidad, longitud y recortar los adaptadores se utilizaron los programas Basic GS Jr Processing HVR1 v1.0 (McElroy, Zagordi, Bull, Luciani y Beerenwinkel, 2013) y PRINSEQ (Huse, Huber, Morrison, Sogin y Welch, 2007), con los archivos FastQ se realizó el mapeo de secuencias contra un genoma de referencia con las plataformas Galaxy (Afgan et al., 2016; Giardine et al., 2005), siguiendo el algoritmo se utilizó la herramienta SAMtools que convierte los archivos SAM a BAM, para indexar los alineamientos los formatos se transformaron a BAI; por último con los archivos FASTA fueron analizados con el programa MEGA 6.0.6 para crear árboles filogenéticos con los que se determinaron las distancias nucleotídicas entre las diferentes muestras.

5. Resultados

5.1 Descripción del tipo de respuesta al tratamiento en pacientes infectados con HCV sometidos a terapia triple.

Se llevó a cabo la secuenciación por el método de Sanger de las regiones codificantes de dos proteínas que intervienen en la inmunidad innata (TICAM1 y NOD1) en 6 pacientes infectados con HCV sometidos a terapia triple (PEG-IFN, RBV y BOC), 4 considerados clínicamente respondedores (ID3, ID4, ID5 e ID7; figura 14) y 2 no respondedores al tratamiento (ID1, ID2; figura 15).

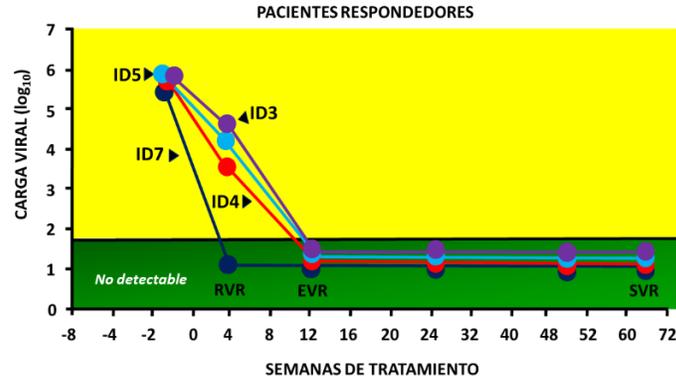


Figura 14. Cargas virales (\log_{10}) de los 4 pacientes Respondedores. Un paciente con RVR (ID7) y 3 con EVR (ID3, ID4, ID5); los 4 pacientes logran una SVR.

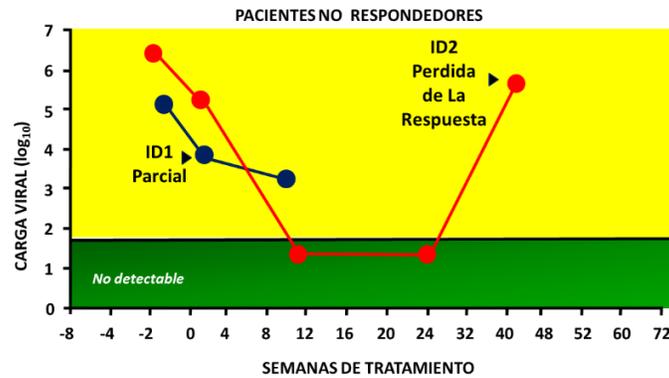


Figura 15. Cargas virales (\log_{10}) de los 2 pacientes No Respondedores. ID1 con respuesta parcial y ID2 con perdida de la respuesta.

5.2 Identificación de SNPs presentes en las moléculas de la Inmunidad Innata TICAM1 y NOD1.

5.2.1 Descripción de SNPs en la molécula TICAM1

Para la molécula TICAM1 se diseñaron 6 pares de primers que generaron una cobertura total de la región codificante (Figura 16). El paciente con **ID7** (Respondedor con **RVR**, es decir, una respuesta rápida a las 4 semanas, presentó un genotipo homocigoto (**AA**) en el SNP rs7255265 (Figura 17) que da lugar a un cambio sinónimo en el codón 7 del mRNA en el amplicón de la molécula de TICAM1 (Figura 16). Este paciente presentó también un genotipo homocigoto (**CC**) en la posición 557 del mRNA del rs2292151 que da lugar a un cambio sinónimo con codificación a ácido aspártico (Figura 19 y Figura 20).

Los pacientes respondedores con EVR (respuesta a las 12 semanas) ID3, ID4 y ID5 presentaron el genotipo homocigoto GG en el rs7255265 (Figura 18) al igual que los pacientes No respondedores ID1 y ID2 y fueron heterocigotos CT para el rs2292151 (Figura 19 y Figura 21). En la figura 22 se muestra un modelado *in silico* de la

molécula TICAM1 donde se resalta el aminoácido treonina codificado dentro del SNP rs7255255.

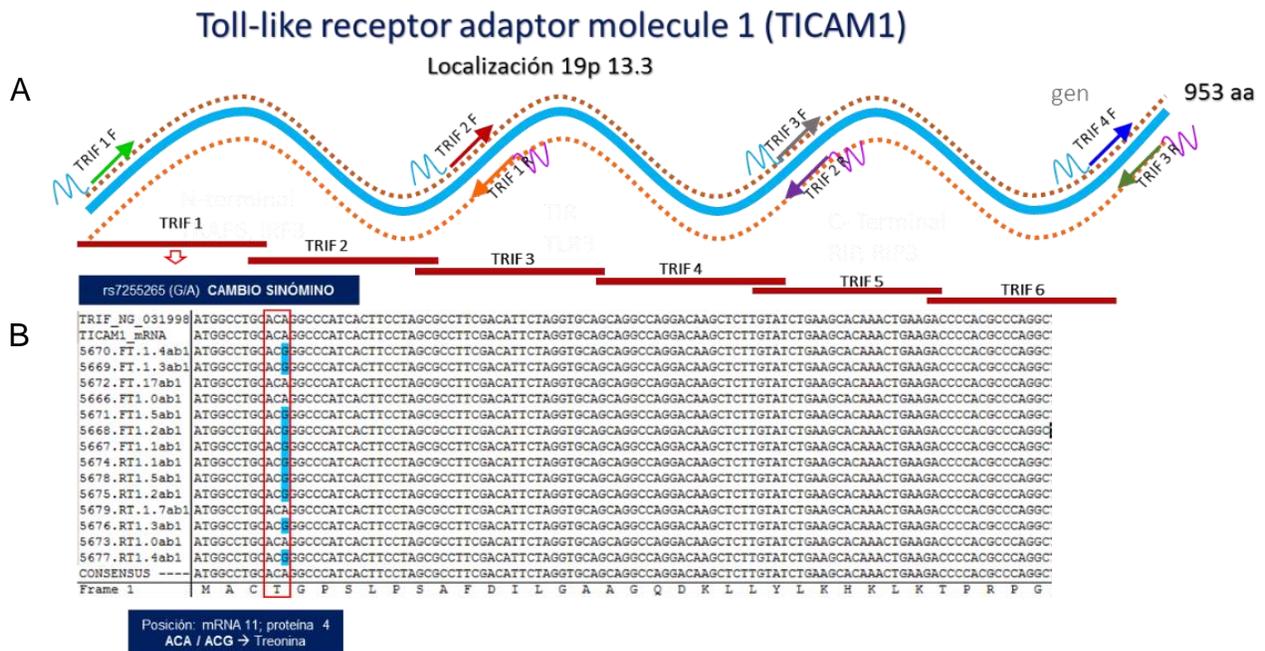


Figura 16. Imagen representativa de la cobertura de la molécula TICAM1. Mediante el diseño de primers (TRIF1- TRIF6) que general amplicones sobrelapados para lograr una cobertura total (A). En el ensamble de secuencias del amplicón TRIF1 (B) se resalta en un recuadro rojo el codón 4 del RNA, el cambio en el SNP rs7255265 G/A es sinónimo. Imágenes obtenidas de GAP v4.11.2.

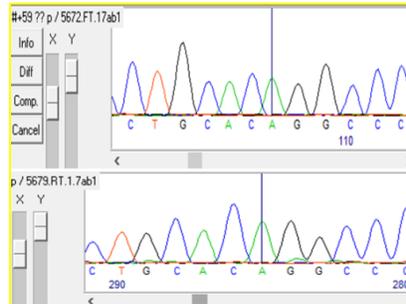


Figura 17. Electroferogramas del primer amplicón de la molécula TRIF. En el que se esquematiza la presencia del Alelo A en el SNP rs7255265 en el paciente ID7.

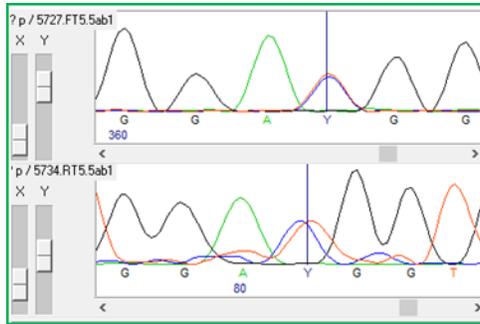


Figura 21. Electroferogramas del quinto amplicón de la molécula TRIF.
 Se observa la presencia de los Alelos C y T en el SNP rs2292151 en el Paciente ID5; de igual manera este genotipo heterocigoto está presente en los pacientes ID 1,2,3,4.

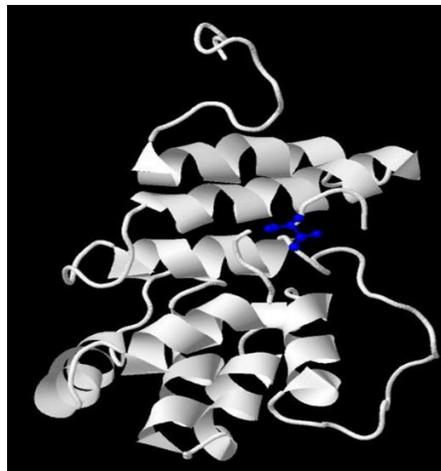


Figura 22. Predicción *in silico* de la molécula TICAM1.

En color azul se señala al aminoácido treonina que se constituye en el cambio sinónimo del SNP rs7255255. Esta región está implicada en la interacción con la molécula de la inmunidad innata TRAF-6
 Imagen obtenida en Jmol.

Para la identificación de SNPs en la molécula NOD1, se diseñaron 3 pares de primers que permiten la amplificación de la región central de la proteína (Figura 23). En el paciente con ID5 (respondedor con EVR, 12 semanas) se identificó el genotipo heterocigoto (**GA**) en el SNP **rs2075820** que da lugar a un cambio no sinónimo en la posición del mRNA 796 (codón 266), este genotipo lleva a la codificación predictiva de cambio de aminoácido de Ácido Glutámico (GAG) a Lisina (AAG) (Figuras 23 y 25); en tanto que en la posición del mRNA 1073 (codón 356) para el SNP **rs376876391** se identificó un genotipo heterocigoto (**GA**), el cual da lugar a un cambio no sinónimo con cambio de arginina (CGC) a histidina (CAC) (Figuras 26 y 28). Los pacientes respondedores con EVR ID3, ID4 y el paciente con RVR ID7, así como los no respondedores ID1 y ID2 presentaron el genotipo homocigoto (GG) tanto en el rs2075820 (Figura 24) como en el rs376876391 (Figura 27) que lleva a la codificación predictiva del aminoácido Ácido Glutámico (rs2075820) y arginina

(rs376876391). También se realizó la predicción *in silico* de la molécula NOD1, destacando la estructura química de los aminoácidos Histidina que representa al SNP rs376887391 y el aminoácido Lisina para el SNP rs2075820 (Figura 29).

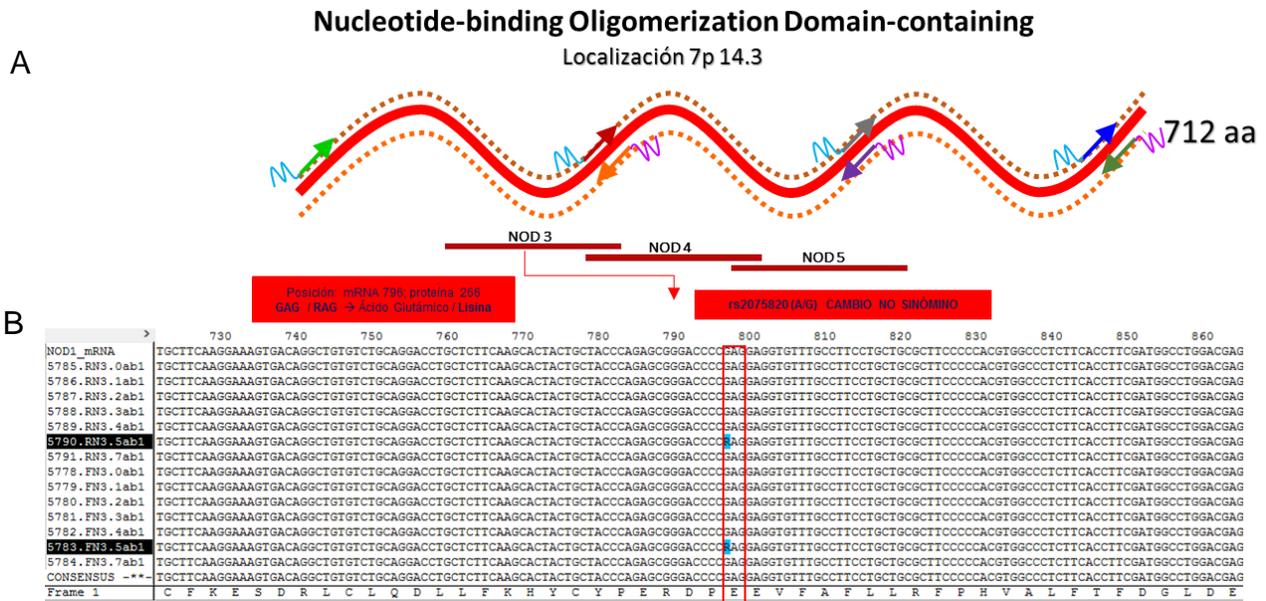


Figura 23. Imagen representativa del diseño de pimer para la secuenciación de la molécula NOD1. Se visualizan los amplicones (NOD3 – NOD6) sobrelapados para lograr una cobertura de la región central de la molécula (A). En el ensamble de Secuencias (B) se resaltan en rojo los codones GAG / RAG (la sigla R indica la presencia de los alelos G y A) en la posición 796 del mRNA, en el que el cambio no sinónimo del rs2075820 da lugar a la codificación del cambio de aminoácido de ácido glutámico a lisina. Imágenes obtenidas de Staden-package Programa GAP v4.11.2.

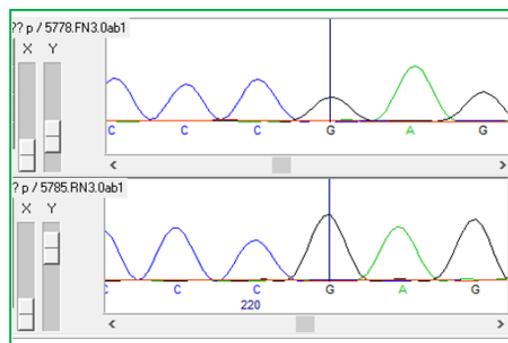


Figura 24. Electroferogramas del tercer amplicón de la molécula NOD1. En el que se observa la presencia del Alelo G en el SNP rs2075820 en la muestra control; de igual manera este genotipo está presente en los pacientes con ID 1,2,3,4,7.

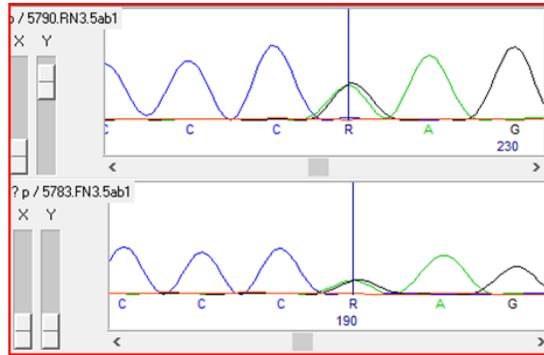


Figura 25. Electroferogramas del tercer amplicón de la molécula NOD1. Se identifica la presencia de los Alelos G y A en el SNP rs2075820 en la molécula NOD1 del paciente con ID5.

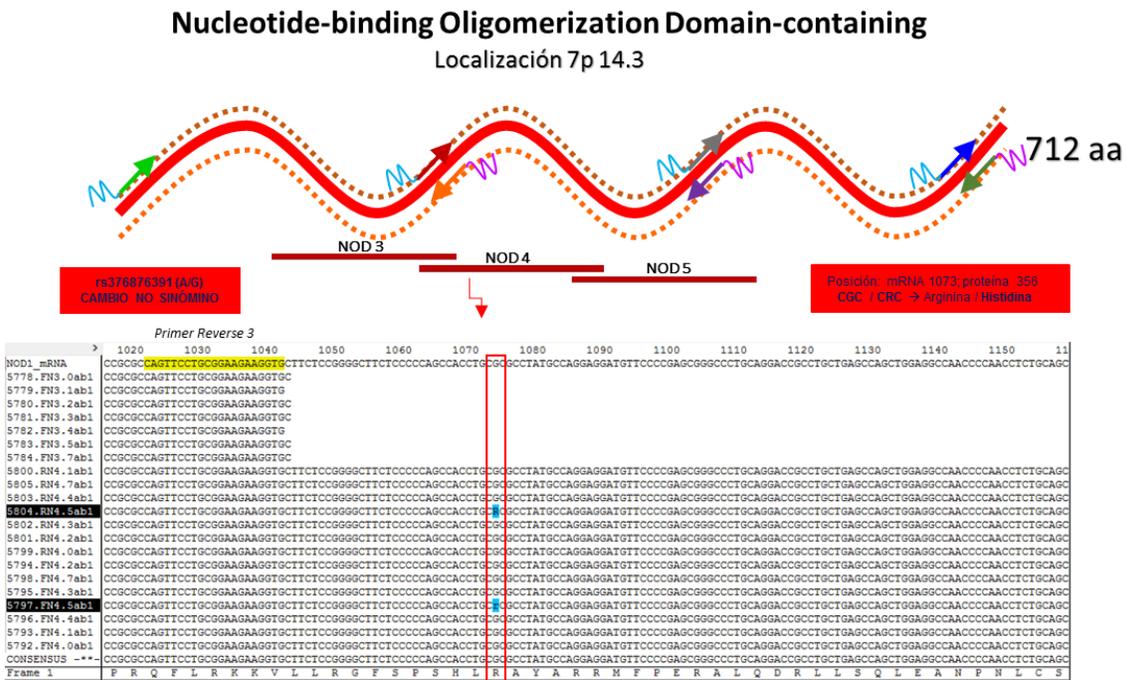


Figura 26. Ensamble de secuencias del amplicón 4 de NOD1. En donde se encuentra el SNP rs376876391 (G/C) en la posición mRNA 1073; se resaltan en rojo los codones CGC / CGR (la sigla R indica la presencia de los alelos G y A), el cambio de G al genotipo heterocigoto (GA) da lugar a la codificación de cambio de aminoácido de Arginina a Histidina presente en el paciente con ID5. Imágenes obtenidas de Staden-package Programa GAP v4.11.2.

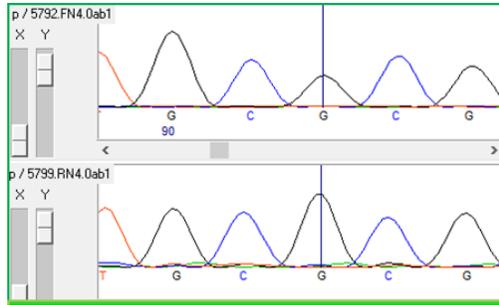


Figura 27. Electroferogramas del cuarto amplicón de la molécula de NOD1. En el que se observa la presencia del alelo G en el SNP rs376876391 en la muestra control; de igual manera este genotipo está presente en los pacientes con ID 1,2,3,4,7.

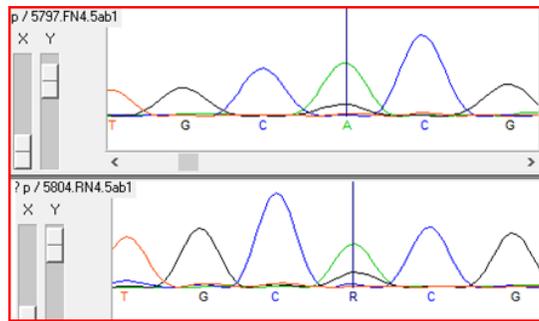


Figura 28. Electroferogramas del cuarto amplicón de la molécula de NOD1. Se esquematiza la presencia de los Alelos A y G en el SNP rs376876391 en la molécula NOD1 Paciente con ID5.



Figura 29. Predicción *in silico* de la molécula NOD1. Destacando en la imagen la estructura química de los aminoácidos Histidina en color amarillo describiendo al SNP rs376876391 codón 356 que codificaría cambio de aminoácido de Arginina a *Histidina*. En color rojo describiendo al SNP rs2075820 posición del codón 266 que codificaría cambio de aminoácido de Ácido Glutámico a *Lisina*, ambos SNP presentes en el paciente ID5. Imagen obtenida por el programa Jmol.

5.2.3 Modelo de Correlación entre los SNPs en TICAM1 y NOD1 en los pacientes con HCV.

Para evaluar si la presencia de los SNPs sinónimos y no sinónimos descritos en este trabajo presentan alguna asociación con la severidad de la enfermedad (descrita como grado de fibrosis), con el tipo de respuesta al tratamiento (RVR, EVR, Nula, Perdida de la respuesta) o con la presencia del genotipo asociado a SVR (CC) en el SNP rs12979860 de IL-28, se realizó un estudio de correlación por el método de Tau B de Kendall. A los pacientes con SNPs en TICAM1 y NOD1 se les asignó un valor de mayor jerarquía en la prueba. Se encontró una correlación positiva moderada (0.65) entre la presencia de polimorfismos en la molécula TICAM1 y una mejor respuesta al tratamiento. De igual forma se encontró una correlación positiva moderada (0.50) entre estos SNPs y una mayor progresión de la enfermedad (mayor grado de fibrosis). Para NOD1 se encontró una correlación negativa baja (-0.39) entre la presencia de SNPs no sinónimos y una mejor respuesta al tratamiento, mientras que para IL-28 no se encontró correlación (Figura 30).

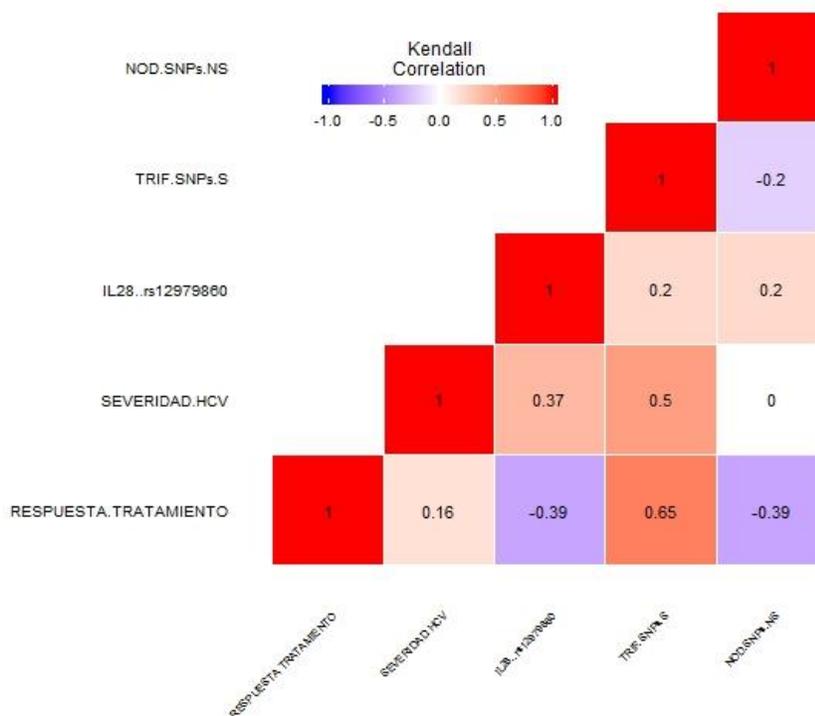


Figura 30. Prueba de correlación Tau B Kendall.

El coeficiente de correlación se representa en colores el rango de rojos una correlación positiva, mientras el rango de azules correlación negativa; o bien sin correlación en color blanco; por lo anterior se observa una correlación positiva moderada (0.65) entre la presencia de los polimorfismos en la molécula TICAM1 y una mejor respuesta al tratamiento, de igual forma se encontró una correlación positiva moderada (0.50) entre los SNPs y una progresión de la enfermedad. Para NOD1 se encontró una correlación negativa baja (-0.39) entre la presencia de SNPs no sinónimos y una mejor respuesta al tratamiento.

5.3 Caracterización de la región HVR1 en pacientes con HCV sometidos a terapia triple.

Para amplificar la región HVR1 se realizaron dos corridas de secuenciación en la plataforma 454 de Roche, posteriormente se realizó el análisis bioinformático para la caracterización de la región HVR1. Se realizaron dos metodologías de análisis, la primera enfocada en calcular las distancias nucleotídicas en muestras basales (cuatro pacientes) y la segunda para realizar un análisis filogenético de la región a lo largo del periodo de tratamiento en dos pacientes. La primera caracterización se realizó a las muestra basales de 4 pacientes, dos respondedores y dos no respondedores (Tabla 4), antes del inicio del tratamiento; se observa una menor distancia nucleotídica entre las variantes virales en los pacientes respondedores (ID3, ID7) en comparación con los pacientes no respondedores (ID1, ID2) (Figura 31). El paciente no respondedor que perdió la respuesta al tratamiento (ID2) presentó la mayor diversidad nucleotídica basal (22.55) seguido por el paciente con respuesta parcial (ID1) con un valor de 20.09, en tanto que los pacientes respondedores presentaron una menor diversidad siendo el paciente con RVR (ID7) el que tuvo la menor diversidad (3.68) y el paciente con EVR (ID3) presento un valor de 16.87 (tabla 4 y figura 31).

PACIENTE	SEMANA DE TRATAMIENTO	CLASIFICACIÓN AL TRATAMIENTO	CARGA VIRAL REPORTADA (LOG)	COMPUTE OVERALL MEAN DISTANCE (π)
ID 1	BASAL	NO RESPONDEDOR	5.4	20.098
ID 2	BASAL	NO RESPONDEDOR	6.36	22.553
ID 3	BASAL	RESPONDEDOR	5.55	16.871
ID 7	BASAL	RESPONDEDOR	5.21	3.685

Tabla 4. Descripción de la distancia media nucleotídica de la región HVR1 en pacientes con HCV en muestra basal.

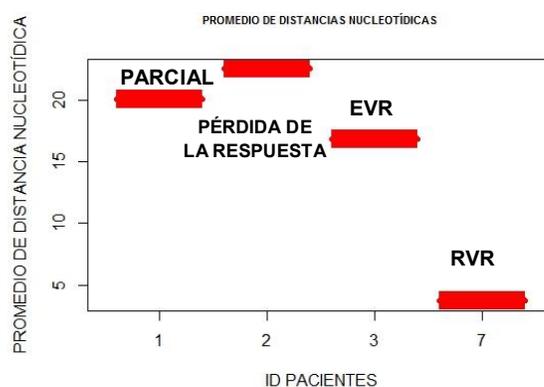


Figura 31. Promedios de las distancias nucleotídicas en la región HVR1 en 4 pacientes infectados con HCV sometidos a terapia triple.

En los pacientes no respondedores, el paciente con ID1 (1) clasificado con una respuesta parcial y ID2 (2) considerado como paciente con pérdida de la respuesta. En los pacientes clasificados como respondedores, el ID3 (3) presento una EVR y el paciente con ID7 una EVR y la menor distancia media nucleotídica.

Para realizar el análisis filogenético de las variantes virales a lo largo del tratamiento, se caracterizaron dos pacientes, un paciente respondedor (ID3) y uno no respondedor (ID2) en los tiempos basal y semana 4 para el ID3 y basal y semanas 4, 12, 24 y 38 para el ID2 como se muestra en la tabla 5.

Con la secuenciación y el posterior análisis bioinformático, se procedió a realizar la reconstrucción filogenética de los pacientes ID3 (Respondedor figura 32) y ID2 (No respondedor figura 33) mediante el método de máxima verosimilitud modelo de kimura 2 parameter model y un bootstrap = 1000.

En el paciente respondedor, se identifican dos poblaciones virales con una diversificación baja entre la muestra basal y la de 4 semanas, estas variantes virales son ya indetectadas en la carga viral de la semana 12; mientras que en el paciente no respondedor se observa mayor diversidad entre las poblaciones virales basales y las presentes en la semana 4 y las poblaciones virales en la semana 24 y 38 son diferentes a las iniciales; un dato que es importante mencionar, es que se logró caracterizar la región HVR1 en este paciente aun cuando los valores de la carga viral a la semana 24 no son detectados por la prueba diagnóstica de carga viral.

PACIENTE	SEMANA DE TRATAMIENTO	CLASIFICACIÓN AL TRATAMIENTO	CARGA VIRAL REPORTADA (LOG)	FECHA TOMA DE MUESTRA	SECUENCIACIÓN HVR1
ID 3	BASAL	RESPONDEDOR	5.55	08-17-12	PRIMERA CORRIDA DE SECUENCIACIÓN
ID 3	SEMANA 4	RESPONDEDOR	4.5	09-06-12	PRIMERA CORRIDA DE SECUENCIACIÓN
ID 2	BASAL	NO RESPONDEDOR	6.36	17-08-2012	SEGUNDA CORRIDA DE SECUENCIACIÓN
ID 2	SEMANA 4	NO RESPONDEDOR	4.72	20-09-2012	SEGUNDA CORRIDA DE SECUENCIACIÓN
ID 2	SEMANA 24	NO RESPONDEDOR	<1.65	03-07-2013	PRIMERA CORRIDA DE SECUENCIACIÓN
ID 2	SEMANA 38	NO RESPONDEDOR	5.48	02-04-2013	PRIMERA CORRIDA DE SECUENCIACIÓN

Tabla 5. Descripción de los pacientes para el análisis filogenético de la región HVR1 a lo largo del tratamiento.

En la tabla se observa el ID3 paciente respondedor, ID2 paciente no respondedor.

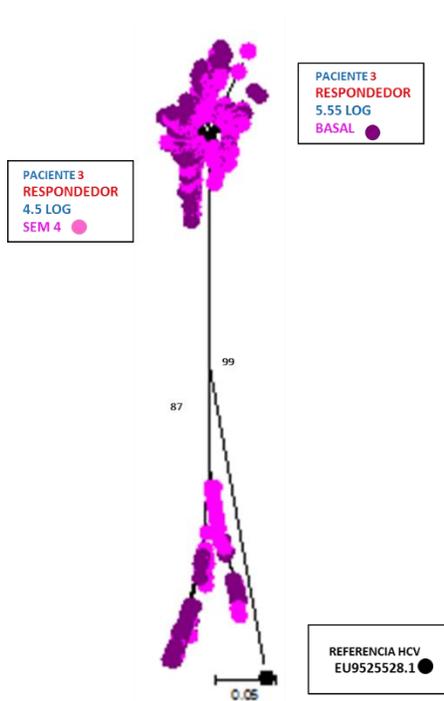


Figura 32. Análisis filogenético por el método de Maximum likelihood.
 Paciente ID3 Respondedor infectado con HCV sometido a terapia triple. Imagen obtenida de MEGA® v 6.0.6

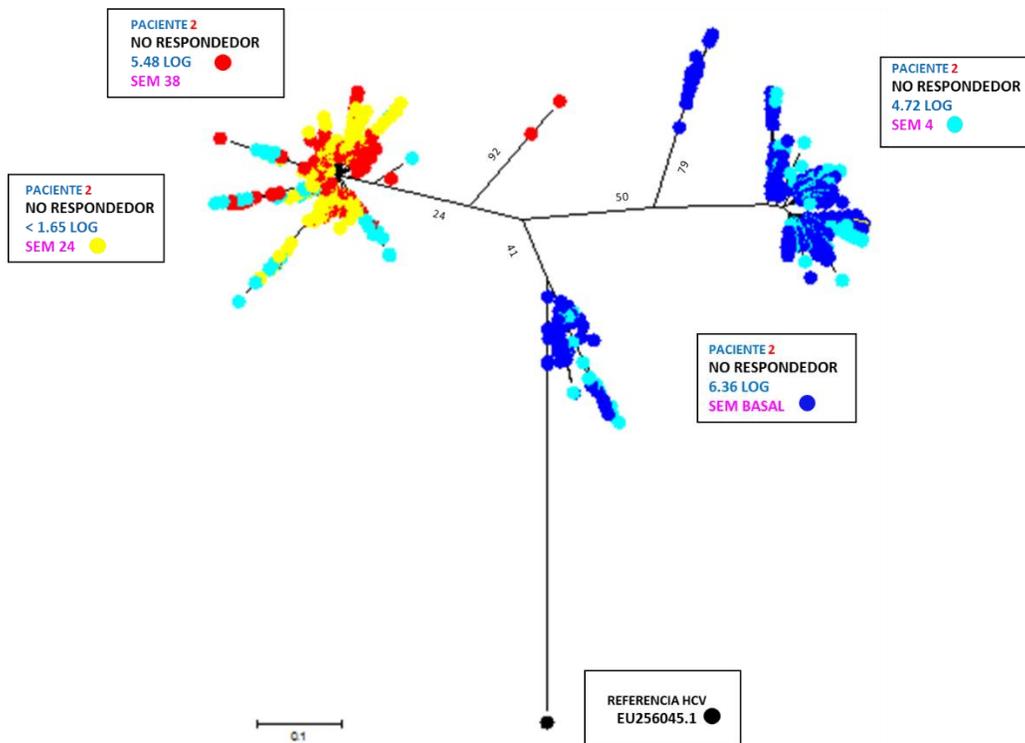


Figura 33. Análisis filogenético por el método de Maximum likelihood.
 Paciente ID2 No Respondedor infectado con HCV sometido a terapia triple. Imagen obtenida de MEGA® v 6.0.

6 Discusión.

La identificación de marcadores moleculares propios del hospedero que presenten asociación con la predicción de la respuesta a la terapia contra HCV o en la progresión de la enfermedad representa uno de los principales objetivos de estudio en hepatitis C; es así que diversos trabajos de investigación se han centrado en estudiar genes candidatos que representen elementos claves en la modulación y función de las moléculas involucradas en el reconocimiento de HCV, entre ellos los genes involucrados en la respuesta inmune innata antiviral (Graells, 2013; Mosbrugger et al., 2010; Sawhney y Visvanathan, 2011).

En el presente trabajo se llevó a cabo la secuenciación por el método de Sanger de las regiones codificantes para dos proteínas que intervienen en la inmunidad innata (TICAM 1 y NOD 1) en 6 pacientes infectados con HCV sometidos a terapia triple (Interferon pegilado, Ribavirina y Boceprevir) con el objetivo de establecer si existe alguna asociación entre los SNPs presentes en estas regiones y el tipo de respuesta al tratamiento, la progresión de la enfermedad o la generación de variantes virales.

Un hallazgo relevante de este proyecto es que, de los pacientes analizados, el único paciente respondedor rápido al tratamiento (RVR) (ID7) (carga viral indetectable a las 4 semanas) posee el genotipo homocigoto AA en el **rs7255265** de TICAM1, el cual está localizado en el codón 7 que codifica para la región proteica del extremo N-terminal, sitio donde la molécula TICAM1 interactúa con la proteína TRAF6 (TNF receptor-associated factor 6) que a su vez induce la vía de señalización de IFN α ; en este mismo paciente, se identificó el genotipo homocigoto CC para el SNP **rs2292151**, el cual está ubicado en el codón 557 en el dominio TIR (Toll/IL-1 receptor) de la proteína en el extremo C-terminal; en dicha región la molécula TICAM1 interacciona con su ligando TLR3, evento importante en el reconocimiento del RNA de doble cadena que se genera durante la replicación viral; tras este reconocimiento, el TLR3 a través de la proteína adaptadora TICAM1 conduce a la activación de diversos factores de transcripción y la translocación al núcleo del IRF3 (factor regulador 3 del interferón) vía que, como ya se mencionó, termina en la activación y producción de IFN α .

En ambos SNPs, a pesar de que los cambios a nivel alélico son sinónimos, pueden existir diferentes repercusiones en la capacidad de expresión de las proteínas y de sus interacciones; así, se ha descrito que la presencia del alelo A en el **rs7255265** aumenta la expresión de IL2 (citocina proinflamatoria) y la presencia del alelo C en el **rs2292151** se ha asociado con el aumento de la producción de TNF α (Ovsyannikova et al., 2011), de tal manera que las variaciones alélicas encontradas en el paciente ID7 podrían estar asociadas a su capacidad de responder de manera temprana (RVR) al tratamiento.

En relación a la asociación entre la molécula TICAM1 y las variantes virales, se ha descrito que la proteasa viral NS3 tiene capacidad de romper a esta proteína adaptadora, evitando la correcta inducción de interferones tipo I y de otras citocinas (Z. Jiang et al., 2004; Ovsyannikova et al., 2011). TICAM1 funciona como proteína adaptadora de unión a TLR3 para mediar la activación del factor de transcripción nuclear IRF 3 (Factor regulador 3 del interferón) a través de TRAF6, y del factor de transcripción NF- KB a través de su unión a NEMO, se ha descrito que mutaciones en TICAM1 en la posición de aminoácidos 250 – 255 impiden la correcta vía de señalización e inducción de estos factores de transcripción (Zhengfan Jiang, Tak W. Mak, Ganes Sen y Xiaoxia Li, 2004); la incisión de la proteína TICAM1 en los aminoácidos Cys 373/ Ser 373 por la proteína NS3 del HCV interrumpe la vía de señalización que conduce a la producción de moléculas proinflamatorias y otras citocinas que desencadenen la respuesta inmune adaptativa, permitiendo una mayor replicación viral en un ambiente de presión selectiva ejercida por la terapia triple que puede dar lugar a la generación de variantes virales (Figura 34) (Zhengfan Jiang et al., 2004).

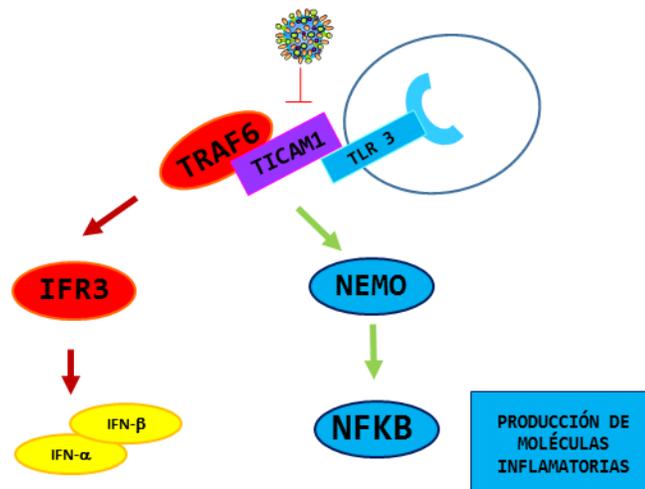


Figura 34. Vía de Señalización mediada por TLR3.

La molécula adaptadora TICAM1 se une a TLR3 y a TRAF6 , el cual induce la producción de IFN α y β y de moléculas inflamatorias. VHC rompe a la molécula TICAM1 en cys-373-ser bloqueando esta vía de señalización.

Hasta el momento no se ha descrito la presencia de SNPs en moléculas de la inmunidad innata en asociación con la infección por HCV, por lo que la descripción de estos SNPs homocigotos en pacientes con RVR debe ser evaluada de manera más amplia, sobre todo tomando en consideración que los pacientes con respuestas menos eficientes presentan genotipos heterocigotos en el **rs2292151** y homocigosis para el alelo ancestral en el **rs7255265**.

En relación a la molécula NOD, se diseñaron 3 pares de primers que permitieron la secuenciación de la región que codifica para el dominio central de la proteína. En el análisis bioinformático se identificaron dos polimorfismos no sinónimos en el paciente respondedor temprano (EVR) ID5; en el primer SNP **rs2075820**, ubicado en el codón 266, se identificó el genotipo GA, en el que la aparición del alelo A (alelo menor) está implicado en un cambio de aminoácido de Ácido Glutámico a **Lisina** según el modelo de predicción de la traducción; para el segundo SNP **rs376876391** (posición 1073 del mRNA), se identificó el genotipo GA en el que, nuevamente, la aparición del alelo A codificaría para un cambio aminoácido de Arginina a **Histidina**.

La región central de la proteína NOD1 está implicada en la oligomerización de la proteína, su correcta activación permite la interacción con la proteína RIPK2, molécula que participa en la modulación de la respuesta inmune innata y adaptativa; además de estar involucrada en el control de la apoptosis (Kutikhin, 2011; Ver Heul, Gakhar, Piper y Subramanian, 2014); aquí se describen SNPs que dan lugar a cambios no sinónimos en esta molécula involucrada en la respuesta ante la infección por HCV y existen pocos estudios de las implicaciones de cambios en esta molécula asociados a procesos patológicos. En este sentido, el alelo A del SNP **rs2075820** se ha asociado con el desarrollo de cáncer y se ha identificado como un marcador de riesgo genómico (genes con un patrón hereditario, cuya alteración aumenta el riesgo a adquirir cierta enfermedad (Klonowska et al., 2015)) y de severidad de la malaria (Kutikhin, 2011; Toure et al., 2012), por su parte, para el SNP **rs376876391** no se ha identificado asociación en ninguna enfermedad hasta el momento.

La proteína NOD1 es una proteína citosólica, de la familia de los receptores NLR que son sensores intracelulares y funcionan como receptores de reconocimiento de patrones tanto de daño como de patógenos (PRR); contiene un dominio caspasa (CARD) en el extremo N terminal, un dominio central de unión a nucleótidos (NBD) y 10 repeticiones de leucinas (LRRs) en su extremo C terminal. NOD1 induce la activación de NF- kappa B a través de RIPK2 (Figura 35), para su correcta función de activación requiere de un proceso de oligomerización mediado por la región central (se ha descrito que la presencia de mutaciones en la región central pueden afectar el proceso de oligomerización necesario para continuar con la vía de señalización que termina en la producción de citocinas pro inflamatorias (Cui, Chen, Wang y Wang, 2014). La presencia de los dos polimorfismos no sinónimos podrían modificar la estructura de la proteína y tener implicaciones en la correcta oligomerización de NOD, ya sea impidiéndola o aumentándola con la consecuente disminución o activación de las vías de señalización mediadas por NOD1 ante las señales de daño generadas por la infección con HCV. En este sentido, solo existen dos reportes de la molécula de NOD1 en asociación con HCV: el primero describe

que en biopsias de pacientes infectados con HCV y analizados por técnicas histoquímicas, se observa una mayor expresión de NOD1 en comparación con tejido sin infección (Mozer-Lisewska et al., 2011); en el segundo se describe la activación y reconocimiento de NOD1 por RNA de doble cadena y su importancia en el proceso de replicación de HCV (Vegna et al., 2016); es así que se requiere continuar con el estudio de las implicaciones que puede tener la presencia de cambios no sinónimos en esta molécula y sus implicaciones en la infección por HCV, ya que con los resultados aquí obtenidos no es posible establecer asociaciones ya que estos cambios se observaron solo en un paciente respondedor.

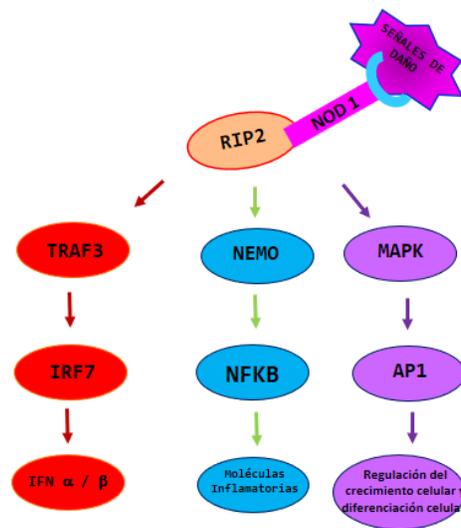


Figura 35. Vía de Señalización de la molécula NOD1.

Tras el reconocimiento de PAMP's (señales de daño) se induce la oligomerización y activación de diversas vías que promueven la producción de IFN α y β , moléculas inflamatorias y moléculas importantes en el crecimiento y diferenciación celular.

La descripción de los SNPs en moléculas de la inmunidad innata implicados en el reconocimiento de HCV se inserta en el mecanismo general de la respuesta inmune ante este agente, el cual depende del tipo de interacciones entre el virus y estas moléculas que, en conjunto, determinan que el virus sea eliminado o se establezca de forma crónica (Heim, 2013; Ver Heul et al., 2014; Zuniga et al., 2015). Las células de la respuesta inmune innata son capaces de reconocer una amplia gama de productos virales a través de sus PRRs como los receptores tipo Toll (TLRs), las helicasas RIG 1 y MDA5, los receptores tipo NOD (NLRs), capaces de reconocer señales de daño causados por el estrés celular o componentes del patógeno (Mozer-Lisewska et al., 2011); si el reconocimiento es adecuado, se da una correcta activación que promueve la detección viral rápida y liberación de citocinas antivirales claves, cuya respuesta puede contener la replicación viral. En particular, la vía de

señalización final de este reconocimiento promueve la producción de interferones de tipo I (IFN I), que pertenecen a un conjunto de citocinas pleiotrópicas que se activan de forma autocrina o paracrina, la unión del IFN a sus receptores induce la vía JAK- STAT cuyo producto final es la activación de un conjunto muy grande de genes denominados genes estimulados por IFN (ISG) que promueven un estado antiviral y activan múltiples células inmunes (Bartenschlager et al., 2013; Heim, 2013). Colectivamente, estos genes facilitan tanto la eliminación del virus de las células infectadas como la protección a las células no infectadas vecinas de la progenie viral entrante. Se ha descrito que en pacientes que responden favorablemente al tratamiento (PEG- IFN y RBV) se induce un aumento de la expresión de un amplio número de genes que participan en la activación de la vía del interferón como ISG15, IFI6, IFI44L, IFI27, MX1, OASL, IRF7, IFIT3, IFITM1, EIF2AK2, HERC5 y APOBEC3A; también se ha observado una sobreexpresión de otros genes involucrados en procesos inmunes (EPSTI1), en adhesión celular (NEXN), en la regulación de la transcripción (HES4) y en otras funciones (GYPB y MT2A). Por el contrario, en pacientes que no responden favorablemente al tratamiento se induce la sobreexpresión de un menor número de ISGs (ISG15, IFI44L, IFI27, IRF7); además, en los no respondedores se observa también un aumento de la expresión de un gen que presenta efecto inhibitorio en la ruta del interferón (USP18) (Figura 36) (Bartenschlager et al., 2013).

El exceso de IFN tipo I promueve la producción de IFN III o IFN λ , los cuales están conformados por tres isoformas que implican a los genes IFN λ 1 (IL29), IFN λ 2 (IL28A) e IFN λ 3 (IL-28B). Los interferones tipo III están relacionados con los miembros de la familia de IL-10 e IL-22 que se ha demostrado que confieren hepatoprotección, además de inducir una cascada de transducción de señales muy parecida a la que inducen los IFNs tipo I, por tanto tienen actividades biológicas similares, ya que ambos tipos de IFNs tienen la habilidad intrínseca de inducir una respuesta antiviral (Ashwin Balagopal, David L. Thomas y Chloe L. Thio, 2010).

De esta forma, y dado el antecedente de asociación entre los SNPs presentes en IL-28 B y la respuesta al tratamiento, se hace pertinente continuar con la búsqueda de marcadores moleculares de respuesta inmune innata y su asociación con el tipo de respuesta con los nuevos antivirales de acción directa así como sus interacciones con las variantes virales.

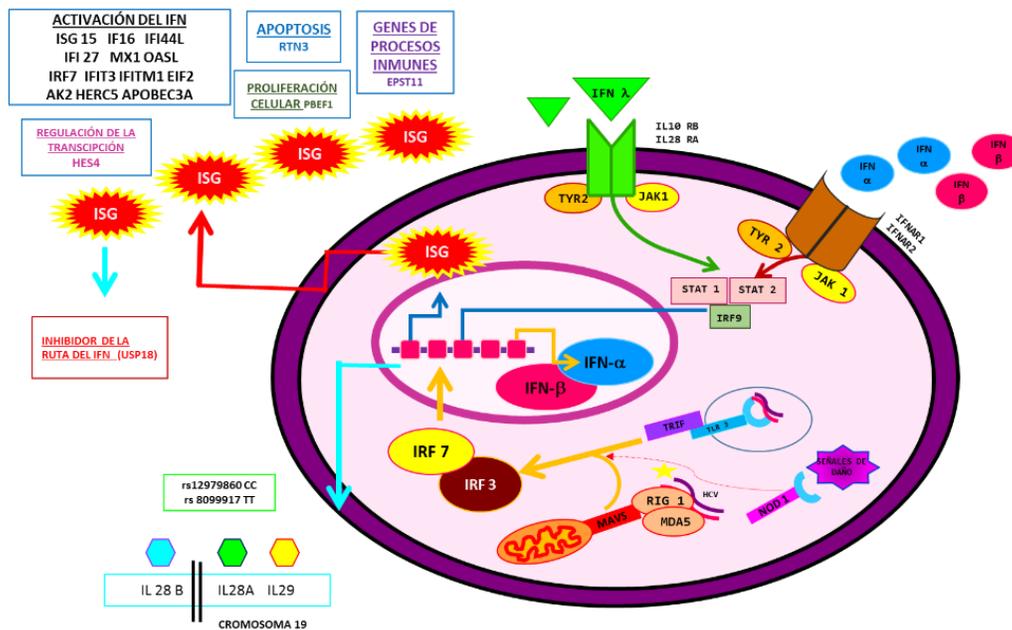


Figura 36. Descripción general de la respuesta inmune innata antiviral para HCV. Adaptado de (Bartenschlager et al., 2013).

El siguiente objetivo del trabajo fue valorar el impacto de las variaciones alélicas y genóticas en los SNPs encontrados, con parámetros clínicos; el análisis por el método de Tau B de Kendall describió una correlación positiva moderada entre la presencia de polimorfismos y una mejor respuesta al tratamiento y una correlación positiva moderada entre estos SNPs y un mayor grado de fibrosis para la molécula TICAM1, en tanto que para NOD1 se encontró una correlación negativa baja entre la presencia de SNPs no sinónimos y una mejor respuesta al tratamiento, en tanto que para IL-28 no se encontró correlación (Figura 30). Este estudio debe ser retomado con un mayor número de pacientes que permitan dar un soporte estadístico sólido a estas correlaciones, sin embargo, como ya se mencionó, una de las principales limitantes de este estudio fue el número de pacientes, debido a que la terapia triple no fue incluida en el cuadro básico de medicamentos mexicanos, por lo que sólo los pacientes que se atienden de forma privada que pudieran cubrir el costo del tratamiento y en protocolos con cohortes muy reducidas tuvieron acceso al tratamiento.

La principal limitante que se encontró en este proyecto fue que la calidad de las muestras no fue la óptima para recuperar el RNA viral, por lo que solo se logró amplificar y secuenciar la región de HVR1 para 4 pacientes (muestras basales) y solo 2 pacientes durante el tratamiento, es muy importante decir que estas muestras se recolectaron en los años 2012 – 2013 y se mantuvieron a bajas temperaturas (-70°), sin embargo, y a pesar de que se realizaron diversas rondas de extracción y

generación de bibliotecas, la viabilidad de las muestras se vio afectada por lo que no se logró identificar en todos los pacientes la región HVR1.

Los pacientes en los que se logró la secuenciación fueron el ID1, ID2, ID3 e ID7 para muestra basal y ID2 y ID3 se lograron secuenciar muestras a lo largo de todo el curso del tratamiento utilizando la amplificación de la región HVR1 de HCV, la cual es representativa de la presión selectiva ejercida en todo el genoma viral por los fármacos y por parte del sistema inmune (Cortes et al., 2014).

El hallazgo más importante al hacer el cálculo del promedio de las distancias nucleotídicas (π) de las 4 muestras basales (figura 31), es que el paciente que inicialmente respondió y al final del tratamiento tuvo una pérdida de la respuesta (ID2), presenta la mayor diversidad basal (22.55), en tanto que el paciente con la respuesta de eliminación viral más eficiente (ID7) es el que presenta la menor diversidad (3.68); aunado a lo anterior, el paciente no respondedor que presenta respuesta parcial, tiene una diversidad nucleotídica elevada (20.09), pero menor que el ID2, y el paciente respondedor a las 12 semanas (ID3) presenta una diversidad menor que los no respondedores (16.87). Lo anterior sugiere que, la existencia de heterogeneidad genética entre los pacientes antes de iniciar el tratamiento, medida como π , pudiera tener utilidad como marcador biológico de predicción al éxito del tratamiento antiviral.

Hasta el momento no se ha utilizado a π como marcador de asociación entre las cuasiespecies presentes antes del tratamiento y el éxito del mismo. Sin embargo, otros parámetros de asociación genética como la entropía de Shannon, el parámetro de Hamming, la distancia genética, la diversidad nucleotídica por sitio y el número de sustituciones genéticas, han sido evaluados con resultados en los que se ha encontrado asociación y en los que no se ha podido establecer (Cortes et al., 2014; Moreau, Levis, Crosbie, Kenny-Walsh y Fanning, 2008).

La existencia de una menor diversidad nucleotídica basal en el paciente ID7 y la presencia de diferencias en los SNPs de respuesta inmune innata, con relación a pacientes con menor eficiencia en la respuesta al tratamiento con terapia triple, sugieren que puede existir una relación entre la identificación de SNP específicos y el tipo de respuesta en función de las poblaciones virales infectantes. La generación de modelos que involucren a los tres parámetros se presenta como una perspectiva real ante los nuevos tratamientos.

Con los avances de la segunda generación de secuenciación (NGS), se han descrito diversos estudios que caracterizan con una gran cobertura diversas regiones que codifican para el virus de la HCV, lo que ha permitido la descripción de diversas asociaciones, entre ellas, con la severidad de la infección, con la transmisión

vertical, con la coinfección entre HCV y otros virus (HIV, HVB) y por último con el tipo de respuesta al tratamiento (Cruz-Rivera et al., 2013; Escobar-Gutierrez et al., 2012; Preciado et al., 2014).

La mayoría de los estudios recientes de NGS se centran principalmente en el análisis de la diversidad en la proteína E2 de HCV. La región HVR1 contiene la zona con mayor variabilidad genética en el genoma de HCV la cual se presenta como un mecanismo que la generar diversidad permite el escape de la respuesta inmunológica mediada por las células T; en este sentido, los hallazgos recientes describen que la pérdida de respuesta de linfocitos T se asocia con una mayor severidad de la infección y tiene implicaciones en el desarrollo de carcinogénesis en pacientes con HCV y como ya se mencionó, se busca el establecimiento entre diversos parámetros de diversidad en la región y el tipo de respuesta con la terapias actuales (Di Lello, Culasso y Campos, 2015).

Es importante continuar con la identificación de marcadores moleculares que ayuden a definir los criterios de predicción a los nuevos protocolos contra HCV; este trabajo describe por primera vez polimorfismos en las moléculas de la inmunidad innata TICAM1 y NOD1 en pacientes infectados con HCV sometidos a terapia triple (PEG-IFN, RBV, BOC), además de evaluar la heterogeneidad genética en la región HVR1 en muestras basales de 4 pacientes con diferente tipo de respuesta al tratamiento. Los resultados obtenidos, sugieren que los polimorfismos presentes en las moléculas de la inmunidad innata TICAM1 y NOD1 y la complejidad de la población viral de la región HVR1, en pacientes con Hepatitis C podrían tener implicaciones importantes en el tipo de respuesta al tratamiento. Estos cambios afectarían principalmente la vía de señalización del Interferón, principal molécula en la respuesta antiviral.

7 Conclusiones.

- Se describen por primera vez diversos SNPs sinónimos y no sinónimos en las regiones codificantes de las moléculas de la inmunidad innata TICAM1 y NOD1 en población mexicana.
- Se estableció una correlación positiva moderada entre la presencia de polimorfismos en la molécula TICAM1 y una mejor respuesta al tratamiento y entre estos SNPs y una mayor progresión de la enfermedad (mayor grado de fibrosis).
- Se describen diferencias en la diversidad nucleotídica basal entre pacientes con diferentes tipos de respuesta al tratamiento que sugieren un potencial uso del parámetro π como marcador de respuesta al mismo.

8. Perspectivas.

- Continuar con la identificación de SNPs presentes en distintas moléculas de la inmunidad innata en pacientes con Hepatitis C sometidos a terapia con antivirales de acción directa, los inhibidores de polimerasa y proteasa.
- Caracterización de la complejidad y dinámica de las variantes virales intrahospedero de las regiones codificantes de la proteasas y polimerasa viral mediante secuenciación masiva paralela en pacientes infectados con Hepatitis C.
- Establecer un modelo que ayude a predecir una respuesta favorable al tratamiento con los nuevos protocolos contra HCV con marcadores moleculares de variantes virales, los datos clínicos de los pacientes y la identificación de diversos SNPs en moléculas de la inmunidad innata.

9. Literatura Citada.

- Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., et al. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res*, 44(W1), W3-W10. doi:10.1093/nar/gkw343
- Ahmed, A., y Felmlee, D. J. (2015). Mechanisms of Hepatitis C Viral Resistance to Direct Acting Antivirals. *Viruses*, 7(12), 6716-6729. doi:10.3390/v7122968
- Akuta, N., Suzuki, F., Kawamura, Y., Yatsuji, H., Sezaki, H., Suzuki, Y., et al. (2007). Predictive factors of early and sustained responses to peginterferon plus ribavirin combination therapy in Japanese patients infected with hepatitis C virus genotype 1b: amino acid substitutions in the core region and low-density lipoprotein cholesterol levels. *J Hepatol*, 46(3), 403-410. doi:10.1016/j.jhep.2006.09.019
- Andrews, S. (2008). Sequence Assembly Using the Staden Package. Staden Quick Guide. Alike 2.0. . *Babraham Bioinformatics*.
- Angelico, M., Koehler-Horst, B., Piccolo, P., Angelico, F., Gentile, S., Francioso, S., et al. (2008). Peginterferon alpha-2a and ribavirin versus peginterferon alpha-2a monotherapy in early virological responders and peginterferon alpha-2a and ribavirin versus peginterferon alpha-2a, ribavirin and amantadine triple therapy in early virological nonresponders: the SMIEC II trial in naive patients with chronic hepatitis C. *Eur J Gastroenterol Hepatol*, 20(7), 680-687. doi:10.1097/MEG.0b013e3282f5196c
- AppliedBiosystems. (2008). SOLiD™ System Accuracy Retrieved from https://www3.appliedbiosystems.com/cms/groups/mcb_marketing/document/s/generaldocuments/cms_057511.pdf
- Ashfaq, U. A., Javed, T., Rehman, S., Nawaz, Z., y Riazuddin, S. (2011). An overview of HCV molecular biology, replication and immune responses. *Viral J*, 8, 161. doi:10.1186/1743-422X-8-161
- Ayllón, M., Cambra, M., Llave, C., y Moriones, E. (2016). *Enfermedades de plantas causadas por virus y viriones*. . España.

- Balagopal, A., Thomas, D. L., y Thio, C. L. (2010). IL28B and the control of hepatitis C virus infection. *Gastroenterology*, 139(6), 1865-1876. doi:10.1053/j.gastro.2010.10.004
- Balagopal, A., Thomas, D. L., y Thio, C. L. (2010). IL28B and the Control of Hepatitis C Virus Infection. *Gastroenterology*, 139(6), 1865-1876. doi:10.1053/j.gastro.2010.10.004
- Bartenschlager, R., Lohmann, V., y Penin, F. (2013). The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. *Nat Rev Microbiol*, 11(7), 482-496. doi:10.1038/nrmicro3046
- Barth, H., Schafer, C., Adah, M. I., Zhang, F., Linhardt, R. J., Toyoda, H., et al. (2003). Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. *J Biol Chem*, 278(42), 41003-41012. doi:10.1074/jbc.M302267200
- Bentley, D. R. (2006). Whole-genome re-sequencing. *Current Opinion in Genetics & Development*, 16(6), 545-552. doi:<http://dx.doi.org/10.1016/j.gde.2006.10.009>
- Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*, 42(Web Server issue), W252-258. doi:10.1093/nar/gku340
- Bonfield, J. K., Smith, K. f., y Staden, R. (1995). A new DNA sequence assembly program. *Nucleic Acids Research*, 23(24), 4992-4999.
- Bonfield, J. K., y Whitwham, A. (2010). Gap5—editing the billion fragment sequence assembly. *Bioinformatics*, 26(14), 1699-1703. doi:10.1093/bioinformatics/btq268
- Bosques-Padilla, F., Trejo-Estrada, R., Campollo-Rivas, O., Cortez-Hernandez, C., Dehesa-Violante, M., Maldonado-Garza, H., et al. (2003). Peginterferon alfa-2a plus ribavirin for treating chronic hepatitis C virus infection: analysis of Mexican patients included in a multicenter international clinical trial. *Ann Hepatol*, 2(3), 135-139.
- Brown, R. A., Gralewski, J. H., Eid, A. J., Knoll, B. M., Finberg, R. W., y Razonable, R. R. (2010). R753Q Single Nucleotide Polymorphism Impairs Toll-like Receptor 2 Recognition of Hepatitis C Virus Core and Nonstructural 3 Proteins. *Transplantation*, 89(7), 811-815. doi:10.1097/TP.0b013e3181cbac18
- Bukh, J., Miller, R. H., y Purcell, R. H. (1995). Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin Liver Dis*, 15(1), 41-63. doi:10.1055/s-2007-1007262
- Carneiro, M. O., Russ, C., Ross, M. G., Gabriel, S. B., Nusbaum, C., y DePristo, M. A. (2012). Pacific biosciences sequencing technology for genotyping and variation discovery in human data. *BMC Genomics*, 13, 375-375. doi:10.1186/1471-2164-13-375
- Checa, M. (2007). Polimorfismos genéticos: Importancia y aplicaciones. *Revista del Instituto Nacional de Enfermedades Respiratorias*, 20, 213 - 221.
- Cordeiro, N., Taroco, R., y Chiparelli, H. (2008). Virus de las hepatitis. *Temas de Bacteriología y Virología Médica*, 477 - 513.

- Cortes, K. C., Zagordi, O., Perlejewski, K., Laskus, T., Maroszek, K., Bukowska-Osko, I., et al. (2014). Deep sequencing of hepatitis C virus hypervariable region 1 reveals no correlation between genetic heterogeneity and antiviral treatment outcome. *BMC Infect Dis*, *14*, 389. doi:10.1186/1471-2334-14-389
- Cruz-Rivera, M., Carpio-Pedroza, J. C., Escobar-Gutierrez, A., Lozano, D., Vergara-Castaneda, A., Rivera-Osorio, P., et al. (2013). Rapid hepatitis C virus divergence among chronically infected individuals. *J Clin Microbiol*, *51*(2), 629-632. doi:10.1128/JCM.03042-12
- Cui, J., Chen, Y., Wang, H. Y., y Wang, R.-F. (2014). Mechanisms and pathways of innate immune activation and regulation in health and cancer. *Human Vaccines & Immunotherapeutics*, *10*(11), 3270-3285. doi:10.4161/21645515.2014.979640
- De Francesco, R., y Migliaccio, G. (2005). Challenges and successes in developing new therapies for hepatitis C. *Nature*, *436*(7053), 953-960. doi:10.1038/nature04080
- De Necoche, R., y Canul, J. (2004). Secuenciación de ácidos nucleicos. Retrieved from http://www.ibt.unam.mx/computo/pdfs/met/secuenciacion_acidos_nucleicos.pdf
- Di Lello, F. A., Culasso, A. C., y Campos, R. H. (2015). Inter and inpatient evolution of hepatitis C virus. *Ann Hepatol*, *14*(4), 442-449.
- Di Maio, V. C., Cento, V., Mirabelli, C., Artese, A., Costa, G., Alcaro, S., et al. (2014). Hepatitis C virus genetic variability and the presence of NS5B resistance-associated mutations as natural polymorphisms in selected genotypes could affect the response to NS5B inhibitors. *Antimicrob Agents Chemother*, *58*(5), 2781-2797. doi:10.1128/AAC.02386-13
- Diago, M., Planas, R. . (2004). *Manual de la Hepatitis C: aspectos biológicos, clínicos y terapéuticos* (Médica Panamericana ed.).
- Díaz de Leon, F., Santos, E. Millan, V., Perez, H., Torre, C. (2003). La hepatitis C. Jefe de la Coordinación de Cirugía General de HG Gonzalo Castañeda ISSTE. *Fac Med UNAM*, *46*, 202 -204
- Douam, F., Ding, Q., y Ploss, A. (2016). Recent advances in understanding hepatitis C. *F1000Res*, *5*. doi:10.12688/f1000research.7354.1
- Dring, M. M., Morrison, M. H., McSharry, B. P., Guinan, K. J., Hagan, R., Irish, H. C. V. R. C., et al. (2011). Innate immune genes synergize to predict increased risk of chronic disease in hepatitis C virus infection. *Proc Natl Acad Sci U S A*, *108*(14), 5736-5741. doi:10.1073/pnas.1016358108
- Du, L., y Tang, H. (2016). An insight into the molecular characteristics of hepatitis C virus for clinicians. *Saudi Med J*, *37*(5), 483-491. doi:10.15537/smj.2016.5.14178
- Durante-Mangoni, E., Iossa, D., y Malgeri, U. (2011). Early viral kinetics and response to treatment in a hepatitis C virus genotype 5 infected patient. *Clin Pract*, *1*(2), e28. doi:10.4081/cp.2011.e28
- EASL Clinical Practice Guidelines: Management of hepatitis C virus infection. *Journal of Hepatology*, *60*(2), 392-420. doi:10.1016/j.jhep.2013.11.003

- Einav, S., Elazar, M., Danieli, T., y Glenn, J. S. (2004). A nucleotide binding motif in hepatitis C virus (HCV) NS4B mediates HCV RNA replication. *J Virol*, 78(20), 11288-11295. doi:10.1128/JVI.78.20.11288-11295.2004
- Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., et al. (1996). Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *N Engl J Med*, 334(2), 77-81. doi:10.1056/NEJM199601113340203
- Escobar-Gutierrez, A., Vazquez-Pichardo, M., Cruz-Rivera, M., Rivera-Osorio, P., Carpio-Pedroza, J. C., Ruiz-Pacheco, J. A., et al. (2012). Identification of hepatitis C virus transmission using a next-generation sequencing approach. *J Clin Microbiol*, 50. doi:10.1128/jcm.00005-12
- Espinoza, L., Trujillo, M., Martinez, R., Panduro, A., Rivas, A., Fierro, N., et al. (2015). Increase of drug use and genotype 3 in HCV infected patients from Central West and Northeast Mexico. *ANNALS of Hepatology*, 14, 642 - 651.
- Farci, P., Alter, H. J., Wong, D. C., Miller, R. H., Govindarajan, S., Engle, R., et al. (1994). Prevention of hepatitis C virus infection in chimpanzees after antibody-mediated in vitro neutralization. *Proc Natl Acad Sci U S A*, 91(16), 7792-7796.
- Fonseca-Coronado, S., Escobar-Gutierrez, A., Ruiz-Tovar, K., Cruz-Rivera, M. Y., Rivera-Osorio, P., Vazquez-Pichardo, M., et al. (2012). Specific detection of naturally occurring hepatitis C virus mutants with resistance to telaprevir and boceprevir (protease inhibitors) among treatment-naive infected individuals. *J Clin Microbiol*, 50(2), 281-287. doi:10.1128/JCM.05842-11
- Franciscus, A. (2009). Progresion de la enfermedad: ¿Qué es la cirrosis? *HSCP*, 5, 1-3.
- Fusco, D. N., y Chung, R. T. (2012). Novel therapies for hepatitis C: insights from the structure of the virus. *Annu Rev Med*, 63, 373-387. doi:10.1146/annurev-med-042010-085715
- Giardine, B., Riemer, C., Hardison, R. C., Burhans, R., Elnitski, L., Shah, P., et al. (2005). Galaxy: a platform for interactive large-scale genome analysis. *Genome Res*, 15(10), 1451-1455. doi:10.1101/gr.4086505
- Gokhale, N. S., Vazquez, C., y Horner, S. M. (2014). Hepatitis C Virus. Strategies to Evade Antiviral Responses. *Future Virol*, 9(12), 1061-1075. doi:10.2217/fvl.14.89
- Gómez I, Á. M. (2003). Biología y métodos diagnósticos de la hepatitis C. *Biomed*, 253 - 268.
- Graells, M. (2013). Hepatitis C e Interleukina 28B (IL28B). Estudio de coste/beneficio de su determinación y frecuencia de polimorfismos en nuestro entorno
- Servicio de Análisis Clínicos. Hospital General Universitario de Alicante*, 14, 71 - 99.
- Greif, G. (2006). Métodos de secuenciación de ácidos nucleicos. Unidad de biología Molecular. *Institut Pasteur Montevideo*.

- Gretch, D. R., Polyak, S. J., Wilson, J. J., Carithers, R. L., Jr., Perkins, J. D., y Corey, L. (1996). Tracking hepatitis C virus quasispecies major and minor variants in symptomatic and asymptomatic liver transplant recipients. *J Virol*, 70(11), 7622-7631.
- Griffin, S. D., Harvey, R., Clarke, D. S., Barclay, W. S., Harris, M., y Rowlands, D. J. (2004). A conserved basic loop in hepatitis C virus p7 protein is required for amantadine-sensitive ion channel activity in mammalian cells but is dispensable for localization to mitochondria. *J Gen Virol*, 85(Pt 2), 451-461. doi:10.1099/vir.0.19634-0
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98. doi:citeulike-article-id:691774
- Hargreaves, A. D., y Mulley, J. F. (2015). Assessing the utility of the Oxford Nanopore MinION for snake venom gland cDNA sequencing. *PeerJ*, 3, e1441. doi:10.7717/peerj.1441
- Heim, M. H. (2013). Innate immunity and HCV. *J Hepatol*, 58(3), 564-574. doi:10.1016/j.jhep.2012.10.005
- Hernández, J., y Martínez, F. (2013). ¿Qué debo saber sobre el tratamiento de la hepatitis C? *gTt*.
- Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M., y Shimotohno, K. (1991). Gene mapping of the putative structural region of the hepatitis C virus genome by in vitro processing analysis. *Proc Natl Acad Sci U S A*, 88(13), 5547-5551.
- Hise, A. G., Traylor, Z., Hall, N. B., Sutherland, L. J., Dahir, S., Ermler, M. E., et al. (2015). Association of Symptoms and Severity of Rift Valley Fever with Genetic Polymorphisms in Human Innate Immune Pathways. *PLoS Neglected Tropical Diseases*, 9(3), e0003584. doi:10.1371/journal.pntd.0003584
- Huang, L., Hwang, J., Sharma, S. D., Hargittai, M. R., Chen, Y., Arnold, J. J., et al. (2005). Hepatitis C virus nonstructural protein 5A (NS5A) is an RNA-binding protein. *J Biol Chem*, 280(43), 36417-36428. doi:10.1074/jbc.M508175200
- Huang, Y.-F., Chen, S.-C., Chiang, Y.-S., Chen, T.-H., y Chiu, K.-P. (2012). Palindromic sequence impedes sequencing-by-ligation mechanism. *BMC Systems Biology*, 6(Suppl 2), S10-S10. doi:10.1186/1752-0509-6-S2-S10
- Huse, S. M., Huber, J. A., Morrison, H. G., Sogin, M. L., y Welch, D. M. (2007). Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biology*, 8(7), R143. doi:10.1186/gb-2007-8-7-r143
- I Burguete-García, A., J Conde-González, C., y Jiménez-Méndez, R. (2011). Hepatitis C seroprevalence and correlation between viral load and viral genotype among primary care clients in Mexico [Seroprevalencia de hepatitis C y correlación entre la carga viral y el genotipo viral en asistentes al nivel primario de atención en México]. 2011, 53.
- Illumina. (2016). An introduction to next generation sequencing Technology Retrieved from http://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf

- Imran, M., Waheed, Y., Manzoor, S., Bilal, M., Ashraf, W., Ali, M., et al. (2012). Interaction of Hepatitis C virus proteins with pattern recognition receptors. *Virology Journal*, 9, 126-126. doi:10.1186/1743-422X-9-126
- Jiang, Z., Mak, T. W., Sen, G., y Li, X. (2004). Toll-like receptor 3-mediated activation of NF-kappaB and IRF3 diverges at Toll-IL-1 receptor domain-containing adapter inducing IFN-beta. *Proc Natl Acad Sci U S A*, 101(10), 3533-3538. doi:10.1073/pnas.0308496101
- Jiang, Z., Mak, T. W., Sen, G., y Li, X. (2004). Toll-like receptor 3-mediated activation of NF-kB and IRF3 diverges at Toll-IL-1 receptor domain-containing adapter inducing IFN-β. *Proceedings of the National Academy of Sciences of the United States of America*, 101(10), 3533-3538. doi:10.1073/pnas.0308496101
- Jimenez-Mendez, R., y Castaneda-Hernandez, G. (2010). Characteristics of hepatitis C treatment with pegylated interferons and ribavirin. *Ann Hepatol*, 9 Suppl, 61-64.
- Jiménez, E., Gobernado, I., y Sanchez, A. (2012). Secuenciación de genoma completo: un salto cualitativo en los estudios genéticos. *Rev Neurol*, 54(11), 692 - 698.
- Klonowska, K., Czubak, K., Wojciechowska, M., Handschuh, L., Zmienko, A., Figlerowicz, M., et al. (2015). Oncogenomic portals for the visualization and analysis of genome-wide cancer data. *Oncotarget*, 7(1).
- Kutikhin, A. G. (2011). Role of NOD1/CARD4 and NOD2/CARD15 gene polymorphisms in cancer etiology. *Human Immunology*, 72(10), 955-968. doi:<http://dx.doi.org/10.1016/j.humimm.2011.06.003>
- Lam, B. P., Jeffers, T., Younoszai, Z., Fazel, Y., y Younossi, Z. M. (2015). The changing landscape of hepatitis C virus therapy: focus on interferon-free treatment. *Therap Adv Gastroenterol*, 8(5), 298-312. doi:10.1177/1756283X15587481
- Lamerdin, J. E., y Carrano, A. V. (1993). Automated fluorescence-based restriction fragment analysis. *Biotechniques*, 15(2), 294-303.
- Langmead, B., Trapnell, C., Pop, M., y Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10(3), R25. doi:10.1186/gb-2009-10-3-r25
- Larrat, S., Kulkarni, O., Claude, J. B., Beugnot, R., Blum, M. G., Fusillier, K., et al. (2015). Ultradeep pyrosequencing of NS3 to predict response to triple therapy with protease inhibitors in previously treated chronic hepatitis C patients. *J Clin Microbiol*, 53(2), 389-397. doi:10.1128/JCM.02547-14
- Lee, W.-P., Stromberg, M. P., Ward, A., Stewart, C., Garrison, E. P., y Marth, G. T. (2014). MOSAIK: A Hash-Based Algorithm for Accurate Next-Generation Sequencing Short-Read Mapping. *PLoS ONE*, 9(3), e90581. doi:10.1371/journal.pone.0090581
- Li, H., y Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14), 1754-1760. doi:10.1093/bioinformatics/btp324
- Li, H., y Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, 26(5), 589-595. doi:10.1093/bioinformatics/btp698

- Li, K., Foy, E., Ferreon, J. C., Nakamura, M., Ferreon, A. C., Ikeda, M., et al. (2005). Immune evasion by hepatitis C virus NS3/4A protease-mediated cleavage of the Toll-like receptor 3 adaptor protein TRIF. *Proc Natl Acad Sci U S A*, 102(8), 2992-2997. doi:10.1073/pnas.0408824102
- Lindenbach, B. D., y Rice, C. M. (2013). The ins and outs of hepatitis C virus entry and assembly. *Nat Rev Microbiol*, 11(10), 688-700. doi:10.1038/nrmicro3098
- Lozach, P. Y., Lortat-Jacob, H., de Lacroix de Lavalette, A., Staropoli, I., Foug, S., Amara, A., et al. (2003). DC-SIGN and L-SIGN are high affinity binding receptors for hepatitis C virus glycoprotein E2. *J Biol Chem*, 278(22), 20358-20366. doi:10.1074/jbc.M301284200
- Márquez, M., Santoscoy, F., y Montoya, H. (2008). Frecuencia y distribución de genotipos del virus de la hepatitis C en población mexicana seleccionada. *Rev Mex Patol Clin*, 55, 79 -87.
- McElroy, K., Zagordi, O., Bull, R., Luciani, F., y Beerenwinkel, N. (2013). Accurate single nucleotide variant detection in viral populations by combining probabilistic clustering with a statistical test of strand bias. *BMC Genomics*, 14. doi:10.1186/1471-2164-14-501
- Mellor, J., Hawkins, A., y Simmonds, P. (1999). Genotype dependence of hepatitis C virus load measurement in commercially available quantitative assays. *J Clin Microbiol*, 37(8), 2525-2532.
- Mendez, R., Salas, F., López, P., Cisneros, L., y Castañeda, G. (2010). Distribution of HCV genotypes and HCV RNA viral load in different regions of Mexico. *J ANNALS of Hepatology*, 9, 33 -39.
- Metzker, M. L. (2005). Emerging technologies in DNA sequencing. *Genome Res.*, 15, 1767-1776.
- Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nat Rev Genet*, 11(1), 31-46. doi:10.1038/nrg2626
- Moradpour, D., Penin, F., y Rice, C. M. (2007). Replication of hepatitis C virus. *Nat Rev Microbiol*, 5(6), 453-463. doi:10.1038/nrmicro1645
- Moreau, I., Levis, J., Crosbie, O., Kenny-Walsh, E., y Fanning, L. J. (2008). Correlation between pre-treatment quasispecies complexity and treatment outcome in chronic HCV genotype 3a. *Virology Journal*, 5, 78-78. doi:10.1186/1743-422X-5-78
- Mosbrugger, T. L., Duggal, P., Goedert, J. J., Kirk, G. D., Hoots, W. K., Tobler, L. H., et al. (2010). Large-scale candidate gene analysis of spontaneous clearance of hepatitis C virus. *J Infect Dis*, 201(9), 1371-1380. doi:10.1086/651606
- Mozer-Lisewska, I., Kowala-Piaskowska, A., Mania, A., Jenek, R., Samara, H., Kaczmarek, E., et al. (2011). Expression of pattern recognition receptors in liver biopsy specimens of children chronically infected with HBV and HCV. *Folia Histochem Cytobiol*, 49(3), 410-416.
- Naito, M., Hayashi, N., Moribe, T., Hagiwara, H., Mita, E., Kanazawa, Y., et al. (1995). Hepatitis C viral quasispecies in hepatitis C virus carriers with normal liver enzymes and patients with type C chronic liver disease. *Hepatology*, 22(2), 407-412.

- Neumann, A. U., Lam, N. P., Dahari, H., Gretch, D. R., Wiley, T. E., Layden, T. J., et al. (1998). Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science*, 282(5386), 103-107.
- Ovsyannikova, I. G., Haralambieva, I. H., Vierkant, R. A., Pankratz, V. S., Jacobson, R. M., y Poland, G. A. (2011). The role of polymorphisms in Toll-like receptors and their associated intracellular signaling genes in measles vaccine immunity. *Hum Genet*, 130(4), 547-561. doi:10.1007/s00439-011-0977-x
- Pallaoro, M., Lahm, A., Biasiol, G., Brunetti, M., Nardella, C., Orsatti, L., et al. (2001). Characterization of the hepatitis C virus NS2/3 processing reaction by using a purified precursor protein. *J Virol*, 75(20), 9939-9946. doi:10.1128/JVI.75.20.9939-9946.2001
- Panduro, A., Melendez, G. E., Fierro, N. A., Madrigal, B. R., Zepeda-Carrillo, E. A., y Román, S. (2011). Epidemiología de las hepatitis virales en México. *Salud Pública de México*, 53, S37-S45.
- Park, C.-W., Cho, M.-C., Hwang, K., Ko, S.-Y., Oh, H.-B., y Lee, H. C. (2014). Comparison of Quasispecies Diversity of HCV between Chronic Hepatitis C and Hepatocellular Carcinoma by Ultradeep Pyrosequencing. *BioMed Research International*, 2014, 853076. doi:10.1155/2014/853076
- Pawlotsky, J.-M. (2016). Hepatitis C Virus Resistance to Direct-Acting Antiviral Drugs in Interferon-Free Regimens. *Gastroenterology*, 151(1), 70-86. doi:<http://dx.doi.org/10.1053/j.gastro.2016.04.003>
- Preciado, M. V., Valva, P., Escobar-Gutierrez, A., Rahal, P., Ruiz-Tovar, K., Yamasaki, L., et al. (2014). Hepatitis C virus molecular evolution: transmission, disease progression and antiviral therapy. *World J Gastroenterol*, 20(43), 15992-16013. doi:10.3748/wjg.v20.i43.15992
- PyMOL. (2012). PyMOL: an Open-Source Molecular graphics tool
- Qian, C., Camps, J., Maluenda, M. D., Civeira, M. P., y Prieto, J. (1992). Replication of hepatitis C virus in peripheral blood mononuclear cells. Effect of alpha-interferon therapy. *J Hepatol*, 16(3), 380-383.
- Ramírez, J., Vargas, G., Tovilla, C., y Fragoso, J. (2013). Polimorfismos de un solo nucleótido (SNP): implicaciones funcionales de los SNP reguladores (rSNP) y de los SNP-ARN estructurales (srSNP) en enfermedades complejas. *Gaceta Médica de México*, 149, 220 - 228.
- Ripoli, M., y Pazienza, V. (2011). Impact of HCV genetic differences on pathobiology of disease. *Expert Rev Anti Infect Ther*, 9(9), 747-759. doi:10.1586/eri.11.94
- Rozo Murillo, L. V., Meneses Escobar, C. A., y Franco Soto, J. (2011). Tecnologías bioinformáticas para el análisis de secuencias de ADN. *Scientia Et Technica*, 116 -121.
- RStudio, T. (2005). RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA.

- Saeed, U., Mazoor, S., Jalal, N., y Zahid Piracha, Z. (2015). Contemplating the Importance of Toll-like Receptors I and II Regarding Human Viral Pathogenesis. *Jundishapur J Microbiol*, 8(1), e13348. doi:10.5812/jjm.13348
- Sakai, A., Claire, M. S., Faulk, K., Govindarajan, S., Emerson, S. U., Purcell, R. H., et al. (2003). The p7 polypeptide of hepatitis C virus is critical for infectivity and contains functionally important genotype-specific sequences. *Proc Natl Acad Sci U S A*, 100(20), 11646-11651. doi:10.1073/pnas.1834545100
- Sanchez-Avila, J. F., Gonzalez, E., Vazquez, V., Suarez, S., y Uribe, M. (2007). Geographical distribution of HCV genotypes in Mexico. *Ann Hepatol*, 6(3), 156-160.
- Sandres, K., Dubois, M., Pasquier, C., Payen, J. L., Alric, L., Duffaut, M., et al. (2000). Genetic heterogeneity of hypervariable region 1 of the hepatitis C virus (HCV) genome and sensitivity of HCV to alpha interferon therapy. *J Virol*, 74(2), 661-668.
- Sanger, F., Nicklen, S., y Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 74(12), 5463-5467.
- Sarrazin, C., Herrmann, E., Bruch, K., y Zeuzem, S. (2002). Hepatitis C virus nonstructural 5A protein and interferon resistance: a new model for testing the reliability of mutational analyses. *J Virol*, 76(21), 11079-11090.
- Sarrazin, C., y Zeuzem, S. (2010). Resistance to direct antiviral agents in patients with hepatitis C virus infection. *Gastroenterology*, 138(2), 447-462. doi:10.1053/j.gastro.2009.11.055
- Sawhney, R., y Visvanathan, K. (2011). Polymorphisms of toll-like receptors and their pathways in viral hepatitis. *Antivir Ther*, 16(4), 443-458. doi:10.3851/IMP1820
- Scarselli, E., Ansuini, H., Cerino, R., Roccasecca, R. M., Acali, S., Filocamo, G., et al. (2002). The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J*, 21(19), 5017-5025.
- Schijman, A., Colina, R., Mukomolov, S., Kalinina, O., García, L., Broor, S., et al. (2004). Comparison of Hepatitis C Viral Loads in Patients with or without Coinfection with Different Genotypes. *Clinical and Diagnostic Laboratory Immunology*, 11(2), 433-435. doi:10.1128/CDLI.11.2.433-435.2004
- Shimizu, Y. K., Hijikata, M., Iwamoto, A., Alter, H. J., Purcell, R. H., y Yoshikura, H. (1994). Neutralizing antibodies against hepatitis C virus and the emergence of neutralization escape mutant viruses. *J Virol*, 68(3), 1494-1500.
- Simmonds, P., Bukh, J., Combet, C., Deleage, G., Enomoto, N., Feinstone, S., et al. (2005). Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology*, 42(4), 962-973. doi:10.1002/hep.20819
- Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., y Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. *Genome Res*, 19(6), 1117-1123. doi:10.1101/gr.089532.108

- Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., et al. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature*, *321*(6071), 674-679. doi:10.1038/321674a0
- Staden, R., Judge, D. P., y Bonfield, J. K. (2003). Analyzing Sequences Using the Staden Package and EMBOSS. In S. A. Krawetz y D. D. Womble (Eds.), *Introduction to Bioinformatics: A Theoretical and Practical Approach* (pp. 393-410). Totowa, NJ: Humana Press.
- Tanaka, Y., Nishida, N., Sugiyama, M., Kurosaki, M., Matsuura, K., Sakamoto, N., et al. (2009). Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet*, *41*(10), 1105-1109. doi:10.1038/ng.449
- Thomas, D. L., Thio, C. L., Martin, M. P., Qi, Y., Ge, D., O'Huigin, C., et al. (2009). Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature*, *461*(7265), 798-801. doi:10.1038/nature08463
- Thompson, A. J., y McHutchison, J. G. (2009). Antiviral resistance and specifically targeted therapy for HCV (STAT-C). *J Viral Hepat*, *16*(6), 377-387.
- Torres B., Venegas Santos, M., y Brahm, J. (2011). Interferones lambdas (ifn λ) en infección con virus hepatitis C. Disponible en <http://www.repositorio.uchile.cl/handle/2250/129064>.
- Toure, O., Konate, S., Sissoko, S., Niangaly, A., Barry, A., Sall, A. H., et al. (2012). Candidate Polymorphisms and Severe Malaria in a Malian Population. *PLoS ONE*, *7*(9), e43987. doi:10.1371/journal.pone.0043987
- Vegna, S., Gregoire, D., Moreau, M., Lassel, P., Durantel, D., Assenat, E., et al. (2016). NOD1 Participates in the Innate Immune Response Triggered by Hepatitis C Virus Polymerase. *Journal of Virology*, *90*(13), 6022-6035. doi:10.1128/JVI.03230-15
- Ver Heul, A. M., Gakhar, L., Piper, R. C., y Subramanian, R. (2014). Crystal Structure of a Complex of NOD1 CARD and Ubiquitin. *PLoS ONE*, *9*(8), e104017. doi:10.1371/journal.pone.0104017
- Vera de León, L., Juárez, A., Navarro, J., Díaz, M., Mendez, J., . (2005). Panorama epidemiológico y situación de la hepatitis C en México *Rev gastroenterol*, *70*, 25 -32.
- Vilchez, G., y Alonso, G. (2009). Alcances y limitaciones de los métodos de epidemiología molecular basados en el análisis de ácidos nucleicos. *Rev. Venezolana de Microbiología*, *29*, 6 -12.
- Wang, G. P., Sherrill-Mix, S. A., Chang, K. M., Quince, C., y Bushman, F. D. (2010). Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. *J Virol*, *84*(12), 6218-6228. doi:10.1128/JVI.02271-09
- Weiner, A. J., Geysen, H. M., Christopherson, C., Hall, J. E., Mason, T. J., Saracco, G., et al. (1992). Evidence for immune selection of hepatitis C virus (HCV) putative envelope glycoprotein variants: potential role in chronic HCV infections. *Proc Natl Acad Sci U S A*, *89*(8), 3468-3472.
- Wilkins, T., Malcolm, J. K., Raina, D., y Schade, R. R. (2010). Hepatitis C: diagnosis and treatment. *Am Fam Physician*, *81*(11), 1351-1357.
- Wilson, M. L., Cai, Y., Hanlon, R., Taylor, S., Chevreux, B., Setubal, J. C., et al. (2013). Sequence verification of synthetic DNA by assembly of

- sequencing reads. *Nucleic Acids Research*, 41(1), e25-e25.
doi:10.1093/nar/gks908
- Workbench, N. S. (2015). NextGen Sequence Workbench.
- Younossi, Z. M., Park, H., Dieterich, D., Saab, S., Ahmed, A., y Gordon, S. C. (2016). Assessment of cost of innovation versus the value of health gains associated with treatment of chronic hepatitis C in the United States: The quality-adjusted cost of care. *Medicine (Baltimore)*, 95(41), e5048.
doi:10.1097/MD.0000000000005048
- Zerbino, D. R. (2010). Using the Velvet de novo assembler for short-read sequencing technologies. *Curr Protoc Bioinformatics*, Chapter 11, Unit 11 15. doi:10.1002/0471250953.bi1105s31
- Zerbino, D. R., y Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, 18(5), 821-829.
doi:10.1101/gr.074492.107
- Zitzer, H., Heilek, G., Truchon, K., Susser, S., Vermehren, J., Sizmann, D., et al. (2013). Second-generation Cobas AmpliPrep/Cobas TaqMan HCV quantitative test for viral load monitoring: a novel dual-probe assay design. *J Clin Microbiol*, 51(2), 571-577. doi:10.1128/JCM.01784-12
- Zuniga, E. I., Macal, M., Lewis, G. M., y Harker, J. A. (2015). Innate and Adaptive Immune Regulation During Chronic Viral Infections. *Annu Rev Virol*, 2(1), 573-597. doi:10.1146/annurev-virology-100114-055226